

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: October 14, 2003, 15:48:26 ; Search time 23109 Seconds
(without alignments)
10933.318 Million cell updates/sec
Title: US-09-754-997A-1
Perfect score: 6176
Sequence: 1 atggcgccggcgacacggg.....taataaactattttccagac 6176
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6176	100.0	6176	6	AX191259	AX191259 Sequence
2	6176	100.0	6176	10	AF176694	AF176694 Mus muscu
3	6038.8	97.8	6219	10	AB052620	AB052620 Mus muscu
4	6025.8	97.6	6222	10	AB052621	AB052621 Mus muscu
5	5810.2	94.1	6301	10	AK122535	AK122535 Mus muscu
6	3756	60.8	3756	6	AX191303	AX191303 Sequence
7	3396.6	55.0	6485	9	AB052622	AB052622 Homo sapi
8	2796	45.3	2796	6	AX191261	AX191261 Sequence
C 9	2743.2	44.4	115135	10	AC112161	AC112161 Mus Muscu
C 10	2743.2	44.4	158055	2	AC110235	AC110235 Mus Muscu
11	2691	43.6	5549	9	AB046848	AB046848 Homo sapi
12	2681.8	43.4	3741	6	AX18406	AX18406 Sequence
C 13	1843.2	29.8	224227	2	AC106509	AC106509 Rattus no
C 14	854.8	13.8	176282	2	AC068507	AC068507 Homo sapi
15	854.8	13.8	198295	9	AC105129	AC105129 Homo sapi
16	825	13.4	825	6	AX191263	AX191263 Sequence
C 17	659.4	10.7	72149	2	AC122776	AC122776 Mus muscu
18	632.2	10.2	2560	9	BC040648	BC040648 Homo sapi
19	571.2	9.2	72149	2	AC122776	AC122776 Mus muscu
C 20	415	6.7	116580	10	AC112162	AC112162 Mus Muscu
21	378.8	6.1	216302	2	AC107847	AC107847 Mus muscu
22	317	5.1	2967	9	BC042054	BC042054 Homo sapi
23	316.6	5.1	2976	6	AR237558	AR237558 Sequence
24	316.6	5.1	3210	6	AR237553	AR237553 Sequence
25	316.6	5.1	3219	6	AR237561	AR237561 Sequence
26	316.6	5.1	3874	6	AR237568	AR237568 Sequence
27	316.4	5.1	2700	6	AX497857	AX497857 Sequence
28	315.4	5.1	3453	6	AR237556	AR237556 Sequence
29	308.4	5.0	3198	10	BC053057	BC053057 Mus muscu
30	306.8	5.0	3146	10	AF026465	AF026465 Mus muscu
31	288	4.7	288	6	AX191277	AX191277 Sequence
32	287.4	4.7	2729	9	AK095529	AK095529 Homo sapi
33	252	4.1	252	6	AX191281	AX191281 Sequence
34	249	4.0	249	6	AX191273	AX191273 Sequence
35	249	4.0	249	6	AX191275	AX191275 Sequence
36	246	4.0	246	6	AX191279	AX191279 Sequence
37	243	3.9	243	6	AX191265	AX191265 Sequence
38	217	3.5	2481	6	AR237560	AR237560 Sequence
39	217	3.5	2715	6	AR237555	AR237555 Sequence
40	217	3.5	2724	6	AR237562	AR237562 Sequence
41	217	3.5	2958	6	AR237557	AR237557 Sequence
C 42	215.8	3.5	147760	9	AC011846	AC011846 Homo sapi
43	215.8	3.5	176282	2	AC068507	AC068507 Homo sapi
44	213.8	3.5	2139	6	AR237563	AR237563 Sequence
45	213.8	3.5	2382	6	AR237566	AR237566 Sequence

ALIGNMENTS

RESULT 1
AX191259
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AX191259
Sequence 1 from Patent WO0149714.
AX191259
GI:15209510

Mus musculus (house mouse)

Mus musculus

Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Salbaum,J.M.
Nope polypeptides, encoding nucleic acids and methods of use
Patent: WO 0149714-A 1 12-JUL-2001;

Salbaum,J.M.

Nope polypeptides, encoding nucleic acids and methods of use

Patent: WO 0149714-A 1 12-JUL-2001;

linear

DNA

6176 bp

Sequence 1 from Patent WO0149714.

PAT 15-AUG-2001

Neurosciences Research Foundation Inc. (US)

FEATURES
Source

Location/Qualifiers

1. 6176

/organism="Mus musculus"

/mol_type="Genomic DNA"

/db_xref="taxon:10090"

1. 3759

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC51120.1"

/db_xref="GI:15209511"

/translation="MARADTGRGLLVLFCLLSARGELPLPQETTVKLSCDEGPLOVI
LGEQAVLDCCTAGATAAGPPTVTMSKGDVLEHENLHLLPNSLSPLEQSDS
DDEALRVTWKTGEGSYSLAHSPLGASQVAVKLAILEDPLSPESQIVBENGAR
RECHTKGAPLITLWEDQVTPPEPRILITLPKLLLOILDVQSDAGSYRVCATNSAR
QFSPSEASITVALRGSLEATRGQDVVIAAPENTTVSGQNVMECVASAPPTPVSW
VRQDKPSTIDVIGRNTLLIASAPRHSVYVCRANKPLTRDPATAAELRLAALP
AISOAPEALYSCAENSATGACAAAPLAVVREGLPSAPTRVATPLSSSVLVAMER
QIGLQADGYTCVAVSAGVQVNDVYQFVNNDDTELQVRLDLEPNTDYEFVYVAYSL
PELHSEQIIGSLHYQKARGVDNVYQFVNNDDTELQVRLDLEPNTDYEFVYVAYSL
CASRTSSPALVHTLDDVPSAAPOLTLSSPNPSDIRVAMLPSSLSNGQVLKYIEVG
LKEQDVSTEVPGNETQITLNSLPNKYVRVIRISAGTAGYGVPSQWQHRTPOVHN
QSHVPFAELKVRAMESLVVSWQPPPHPTQISGYKLYMGVEVGEESADGDRPPGR
GDQAVDQVPLKVKKQYELTQVPRPYEVKLVAFKNHEDGYAAVWKTEKAPTP
DLPIQRGDLPLPAHVAESNSTSIWLRKKPDPFTTVKIVNYTVLRFPGWGLRNASLVT
YVTSGEDILIGLKPFTKYBFAVSHGVDMDPFGSVVERSTLDPDRSTPSPDLRLS
PLTPSTVRLHCPTPENGEIVEYLILYNNHTOPEHOWTLLTTGNIIFSAREVHGLS
DTRYFRMGARTEVGPSPRELQDVITLQETSPDSLDVHVTGIIIVGLGLLCLLAC
MCAGLRQSHREALPGLSSGTPGNPALYTRARLGGPSPVPAHAHELESILPRPQDMS
PPSDEVEDAEHSLMGGSVSDCRGHSKRKISWAQAGGNWAGSWSLSPQSGRPPA
LTRALLPAGTQOTLLOALVYDGIKSNRKKPSACRNOVEAEVIVHSDFGASKGCP
DLHLQDLPEERLPAETLPTSGDAVDLSQADMLGRELGSCQPTTSGPERLTLCLPEAA
SNSCSDLOPSTALTEAPGSKQKALCLPLTVSPSLPRAPVSSAQVP"

BASE COUNT 1316 a 1756 c 1634 g 1470 t

ORIGIN

Query Match 100.0%; Score 6176; DB 6; Length 6176;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCGGCGGACACGCGGCTCTGGTGCTGACCTTCTGCTGCTGCTGCTGCGG 60
DB 1 ATGGCGCGGCGGACACGCGGCTCTGGTGCTGACCTTCTGCTGCTGCTGCTGCGG 60

QY 61 CGCGGGGAGCTGCCATTCGCCACGAGACAACTGTCAAGCTGAGCTGTGATCAGGACCC 120
DB 61 CGCGGGGAGCTGCCATTCGCCACGAGACAACTGTCAAGCTGAGCTGTGATCAGGACCC 120

QY 121 CTGCAAGTGATCTCGGCGCTCGAGCAGGCTGTGGTGCTGCACTGTGCGGCTACA 180
DB 121 CTGCAAGTGATCTCGGCGCTCGAGCAGGCTGTGGTGCTGCACTGTGCGGCTACA 180

QY 181 GTCTGTGGGCTCCGACCAAGGCTGATGAGCAAGGATGGAGACACTGTACTAGAGCAT 240
DB 181 GTCTGTGGGCTCCGACCAAGGCTGATGAGCAAGGATGGAGACACTGTACTAGAGCAT 240

QY 241 GAGAACCTGCACTGCTGACCAATGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 241 GAGAACCTGCACTGCTGACCAATGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

QY 301 GACAGCGATGATGAGGAGCTCTTAGGATCTGGAAGTCACTGAGGCGCAGCTATTCTGT 360
DB 301 GACAGCGATGATGAGGAGCTCTTAGGATCTGGAAGTCACTGAGGCGCAGCTATTCTGT 360

QY 361 CTGGCCACAGCGCGCTAGGAGTGTGGCGCAGCAGGTTGCTGTGTGCTCAAGCTTGCCACA 420
DB 361 CTGGCCACAGCGCGCTAGGAGTGTGGCGCAGCAGGTTGCTGTGTGCTCAAGCTTGCCACA 420

QY 421 CTGGAAGACTTCTCTGACCCCGAGTCCAGATTGTTGGAGAGAGAACGGGACAGCACGC 480
DB 421 CTGGAAGACTTCTCTGACCCCGAGTCCAGATTGTTGGAGAGAGAACGGGACAGCACGC 480

QY 481 TTTGAATGTCACACCAAGGCGCTTCACGCCCCCATCATTTCTTGGGAAAAGGACGAGTG 540

DB 481 TTTGAATGTCACACCAAGGCGCTTCACGCCCCCATCATTTACTTGGGAAAAGGACGAGTG 540

QY 541 ACCGTGCTGAGGAGCGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCACATCCTAGAT 600
DB 541 ACCGTGCTGAGGAGCGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCACATCCTAGAT 600

QY 601 GTCCAGGACAGTGTGATGAGGCTCTCTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAACGA 660
DB 601 GTCCAGGACAGTGTGATGAGGCTCTCTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAACGA 660

QY 661 TTACAGCCAGGAGGCTCGCTCATCTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAAGGG 720
DB 661 TTACAGCCAGGAGGCTCGCTCATCTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAAGGG 720

QY 721 CAGGATGTGCTCATTTGTGGAGCGCCAGAGAACCAACGAGTGTCTGGACAGATGTA 780
DB 721 CAGGATGTGCTCATTTGTGGAGCGCCAGAGAACCAACGAGTGTCTGGACAGATGTA 780

QY 781 GTGATGGAGTGGTGGCTCTGTGACCCCAACCCCTTTTGTGTCTGGGTCCGACGAGT 840
DB 781 GTGATGGAGTGGTGGCTCTGTGACCCCAACCCCTTTTGTGTCTGGGTCCGACGAGT 840

QY 841 GGAAGAGCTATCTCCAGGATGTCTCATGTTCTGGGCGGACCAATCTACTCATCGCCAGC 900
DB 841 GGAAGAGCTATCTCCAGGATGTCTCATGTTCTGGGCGGACCAATCTACTCATCGCCAGC 900

QY 901 CGCAGGCTCGGCTCTGGAGTCTATGTCTCCCGAGGCAACAGAGCCCTCACGGGTGAC 960
DB 901 CGCAGGCTCGGCTCTGGAGTCTATGTCTCCCGAGGCAACAGAGCCCTCACGGGTGAC 960

QY 961 TTGCGCATCTCGGCTGTGAGCTCCGAGTGTCTGTGCGCCAGCCATCTCCGAGGACCC 1020
DB 961 TTGCGCATCTCGGCTGTGAGCTCCGAGTGTCTGTGCGCCAGCCATCTCCGAGGACCC 1020

QY 1021 GAGGCGCTCTCGCGGACCGCGGCGGACGACACCGCGGCTTCGTGTGCGGCGCTCCGGGAG 1080
DB 1021 GAGGCGCTCTCGCGGACCGCGGCGGACGACACCGCGGCTTCGTGTGCGGCGCTCCGGGAG 1080

QY 1081 CCACGGCGCGCTGCACTGCTGCAAGCGGATCCCGTTGCGNACCAATGGGCGCTC 1140
DB 1081 CCACGGCGCGCTGCACTGCTGCAAGCGGATCCCGTTGCGNACCAATGGGCGCTC 1140

QY 1141 AAGTGTGAGGCGCGTGGCGGAGCTTGGTCACTCATGATCGGCTTCAGAGAGCTGCG 1200
DB 1141 AAGTGTGAGGCGCGTGGCGGAGCTTGGTCACTCATGATCGGCTTCAGAGAGCTGCG 1200

QY 1201 TACTACAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
DB 1201 TACTACAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260

QY 1261 GTAGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
DB 1261 GTAGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320

QY 1321 AGCTCTCTGTGCTGTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
DB 1321 AGCTCTCTGTGCTGTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380

QY 1381 TTCTCTCTTCACTACCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
DB 1381 TTCTCTCTTCACTACCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440

QY 1441 AACAAATGACACACAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1441 AACAAATGACACACAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500

QY 1501 TACGTGTGTGCTTACTCTCCAGTGGGCGCAGCGGAACTCCAGCCAGCCCTGCTGAT 1560
DB 1501 TACGTGTGTGCTTACTCTCCAGTGGGCGCAGCGGAACTCCAGCCAGCCCTGCTGAT 1560

QY 1561 ACATCTGGACGATGTCTCCAGCGGACGACCCCAAGCTTACTCTTGTCCAGGCCCCCAACCCCTCG 1620

QY 3781 GGAAATGCGATGCGTACACATGTGTGTCTAGAGATATCCATAGTCTTTGGAGC 3840
DB |||||
3781 GGAAATGCGATGCGTACACATGTGTGTCTAGAGATATCCATAGTCTTTGGAGC 3840
QY |||||
3841 CTCCTTAGGGTCTTTGGGCTGGGGTGGGAGAACTTTTACTCTCCCTCATATTTCTGCATCA 3900
DB |||||
3841 CTCCTTAGGGTCTTTGGGCTGGGGTGGGAGAACTTTTACTCTCCCTCATATTTCTGCATCA 3900
QY |||||
3901 CATACAGGAGGACTTTGAGACACAGCTCTGTGTAAATGGACACAGCTGTGGAAGTCTGTGTGT 3960
DB |||||
3901 CATACAGGAGGACTTTGAGACACAGCTCTGTGTAAATGGACACAGCTGTGGAAGTCTGTGTGT 3960
QY |||||
3961 GT 4020
DB |||||
3961 GT 4020
QY |||||
4021 GCCTAGTTGACCCCTCGTGGCAGGATGGTGTAAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4080
DB |||||
4021 GCCTAGTTGACCCCTCGTGGCAGGATGGTGTAAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4080
QY |||||
4081 TTTAGCCTTGTCACTAGCCTTTTATACACTCTGAGAGTGTCTCCAGTGTGTGTGTGTGTGTGTGT 4140
DB |||||
4081 TTTAGCCTTGTCACTAGCCTTTTATACACTCTGAGAGTGTCTCCAGTGTGTGTGTGTGTGTGTGT 4140
QY |||||
4141 AAAGACAGCGCCAGCCCTCTCTGT 4200
DB |||||
4141 AAAGACAGCGCCAGCCCTCTCTGT 4200
QY |||||
4201 GGGCCTATGACACCGGAGCTTACACAGATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4260
DB |||||
4201 GGGCCTATGACACCGGAGCTTACACAGATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4260
QY |||||
4261 CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTTTGAGAGTCTCTGTGTGTGTGTGTGTGT 4320
DB |||||
4261 CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTTTGAGAGTCTCTGTGTGTGTGTGTGTGT 4320
QY |||||
4321 TTGTAATCTGGGAAGTCTGT 4380
DB |||||
4321 TTGTAATCTGGGAAGTCTGT 4380
QY |||||
4381 CCTTCTCTCATTTGT 4440
DB |||||
4381 CCTTCTCTCATTTGT 4440
QY |||||
4441 CTTCTCTTCCATGATGCGCCCTGATTTCT 4500
DB |||||
4441 CTTCTCTTCCATGATGCGCCCTGATTTCT 4500
QY |||||
4501 CTTGTCTTTATGTCTGT 4560
DB |||||
4501 CTTGTCTTTATGTCTGT 4560
QY |||||
4561 ACATGACTACCTCATGTCTGTCTGT 4620
DB |||||
4561 ACATGACTACCTCATGTCTGTCTGT 4620
QY |||||
4621 CTGCGCAACCGCTTCTCTGGGAGATGAGCCCACTCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4680
DB |||||
4621 CTGCGCAACCGCTTCTCTGGGAGATGAGCCCACTCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4680
QY |||||
4681 TGGCTTGTCTGGTGTCTGCTCCCTCTGT 4740
DB |||||
4681 TGGCTTGTCTGGTGTCTGCTCCCTCTGT 4740
QY |||||
4741 GAGTCAGGCTGCAAGGGGAAACCCCAAGAGGGGTAGGAGTGAAGAACCAAGAGGCTGA 4800
DB |||||
4741 GAGTCAGGCTGCAAGGGGAAACCCCAAGAGGGGTAGGAGTGAAGAACCAAGAGGCTGA 4800
QY |||||
4801 AAAAAATGGTGCACCACTGTGACACAGAGATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4860
DB |||||
4801 AAAAAATGGTGCACCACTGTGACACAGAGATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4860
QY |||||
4861 CTGGCTGAAATGAAAGGGGTGAGGAGGGGAGCTACTGGGGCTGCCATGGCTCAGTTCC 4920

DB 4861 CTGGCTGAAATGAAAGGGGTGAGGAGGGGAGCTACTGGGGCTGCCATGGCTCAGTTCC 4920
QY |||||
4921 CTGACCTTGAGCCCTCAACCTGGCTTTCAGAGTAGCAAGAGTGTCTCTCAAGATGTGTGT 4980
DB |||||
4921 CTGACCTTGAGCCCTCAACCTGGCTTTCAGAGTAGCAAGAGTGTCTCTCAAGATGTGTGT 4980
QY |||||
4981 AAGGGAAGTCTTTGCATAGGAAAGGGCGGCTGGCTCATTTTATTTTATCTTTCTTTTACA 5040
DB |||||
4981 AAGGGAAGTCTTTGCATAGGAAAGGGCGGCTGGCTCATTTTATTTTATCTTTCTTTTACA 5040
QY |||||
5041 CTGAATCCCAAAATCATCTTACCAAAAGGGCCAAAGCCTGACTGTGTATTTCTCTGAGTCAC 5100
DB |||||
5041 CTGAATCCCAAAATCATCTTACCAAAAGGGCCAAAGCCTGACTGTGTATTTCTCTGAGTCAC 5100
QY |||||
5101 AAGAGCCATGCACTCTCTGTGTCTCACCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5160
DB |||||
5101 AAGAGCCATGCACTCTCTGTGTCTCACCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5160
QY |||||
5161 GTGGCATCTGTGTCTTTGTCTGCACATCTTTCTATTTTCAACTGGCTGGGCACATCAAGTGT 5220
DB |||||
5161 GTGGCATCTGTGTCTTTGTCTGCACATCTTTCTATTTTCAACTGGCTGGGCACATCAAGTGT 5220
QY |||||
5221 AACTCTGGCTTCTGGGCCAAGTTAGAAATAACAGTCTATTTTCCCTTTTATTTTATTTT 5280
DB |||||
5221 AACTCTGGCTTCTGGGCCAAGTTAGAAATAACAGTCTATTTTCCCTTTTATTTTATTTT 5280
QY |||||
5281 TTTTATTTTATTTTATGTCTTTTCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5340
DB |||||
5281 TTTTATTTTATTTTATGTCTTTTCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5340
QY |||||
5341 AGCCTTGTGTCTCACTCATGT 5400
DB |||||
5341 AGCCTTGTGTCTCACTCATGT 5400
QY |||||
5401 CAGCCTATGATAACACTAAAGATTTAAATGCTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5460
DB |||||
5401 CAGCCTATGATAACACTAAAGATTTAAATGCTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5460
QY |||||
5461 TGTCACTTGAATCTTCTATAGCATTTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5520
DB |||||
5461 TGTCACTTGAATCTTCTATAGCATTTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5520
QY |||||
5521 CTTTCAATCTCTGAAAGTCAAGAGCTCATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5580
DB |||||
5521 CTTTCAATCTCTGAAAGTCAAGAGCTCATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5580
QY |||||
5581 TTGGCTTGAATAAACCTTTTAAACAGT 5640
DB |||||
5581 TTGGCTTGAATAAACCTTTTAAACAGT 5640
QY |||||
5641 TCCTCATGTCTTGT 5700
DB |||||
5641 TCCTCATGTCTTGT 5700
QY |||||
5701 TGAGACAAATTTATTAACCAATAGGCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5760
DB |||||
5701 TGAGACAAATTTATTAACCAATAGGCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5760
QY |||||
5761 TAAATCTGTGTATCTGCTCTATTAACCTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5820
DB |||||
5761 TAAATCTGTGTATCTGCTCTATTAACCTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5820
QY |||||
5821 TTAATATGATTTGGGGTCTGAGTAACTGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5880
DB |||||
5821 TTAATATGATTTGGGGTCTGAGTAACTGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5880
QY |||||
5881 AGGTCTCTTATAGGAAAGTCTGT 5940
DB |||||
5881 AGGTCTCTTATAGGAAAGTCTGT 5940
QY |||||
5941 ATGGATGCTTAAAGTCTTTTGGAGTCTCATTTTAAAGATTTCTTTTCTCGAAGCATGACAA 6000

Db 5941 ATGATGCTTAAGTCTTTCGAGTCTCATTTAAGAAATTCCTTTTCTCGAAGCATGACAA 6000
Qy 6001 GTGTATCGCAATCTTACATGCTCACTGTTTACCTGCTAGTGTGTGCTGGGTATTT 6060
Db 6001 GTGTATCGCAATCTTACATGCTCACTGTTTACCTGCTAGTGTGTGCTGGGTATTT 6060
Qy 6061 AATTGCACTTTCCAGCATGCTTCCCTCTTACAAATATGATATTTCTTATTGTTACAC 6120
Db 6061 AATTGCACTTTCCAGCATGCTTCCCTCTTACAAATATGATATTTCTTATTGTTACAC 6120
Qy 6121 TAAGTGTGTGATCATGATCTGCTCCCTGTAAGAAATTAATAACTATTTTCCAGAC 6176
Db 6121 TAAGTGTGTGATCATGATCTGCTCCCTGTAAGAAATTAATAACTATTTTCCAGAC 6176

RESULT 2
AF176694 6176 bp mRNA linear ROD 27-APR-2000
LOCUS Mus musculus neighbor of Punc ell protein (Nope) mRNA, complete
cde.
ACCESSION AF176694
VERSION AF176694.1 GI:7650185
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
TITLE Cloning and expression of nope, a new mouse gene of the immunoglobulin superfamily related to guidance receptors
JOURNAL Genomics 64 (1), 15-23 (2000)
MEDLINE 20175427
PubMed 10708514
REFERENCE
AUTHORS Salbaum, J.M.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1999) The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121, USA

FEATURES
source
1. .6176
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB"
/db_xref="taxon:10090"
/chromosome="9"
/map="band D-E1; 24.59 cR from D9Mit143"
1. .6176
/gene="Nope"
1. .3759
/gene="Nope"
/note="immunoglobulin superfamily; related to guidance receptors"
/codon_start=1
/product="neighbor of Punc ell protein"
/protein_id="AAF65930.1"
/db_xref="GI:7650186"
/translation="MARADTGRGLLVITFCLLSARGELPLIPEQTTVKLSDEBPLQVI
LGPQAVVKTGATAGPRTVTWSKGDVLEHNLHLLPNGLSPLSEBDS
DDEALRIWKTGATAGPRTVTWSKGDVLEHNLHLLPNGLSPLSEBDS
FECHTGLPILITWEKQVTPPEPLITLTPKMLLQILDVQSDAGSYRCAVNSAR
QRPSQASLTVALRGLSEATRGQDVIVAAPENTIVVSGQNVMECVASADPTPFVSM
VRQKGPISDVIIVLRTNLLIASAQRHSGVYCRANKPLTRDPATAAAEELRVLAAP
AISOQALSKRSTASTARFVCRASGEPRPALHLDGIPURPNRGRVKVQGGGSLVIT
QIGQDAGYQCAENAGTAAAPLAVVVRGLSPAPRTVATPLSSSLVLAWER
PELHSEQIIGFSLHQKARGVNVYQFVAVNDDTELQVRDLEPNTDYEFTVAYSOL
GASSTSPALVITDDVPSAPOLITLSPNPSDIRVAWLPLPSLSNGVLYKXIYG
LGKEDVPSVEPCNETQLTSLQPNKIVRVISACTGAGYGVPSOMHRTFVGN
QSHVPPAPBLVRKAMESLVSPQPPHPTQISGKLYWGEVETREADGRPPGR
GDQAWDPGLVRKKKXQYELTQLVGRPYEKLVAFNKHEGDAVAAWKGTEKATP
DLPTQRGPPPLPAPHAESNSTSIWLRWKKPFTTVKIVNYTVRFGPWGLRNASLVT
YVTSGGDILIGLKLPTKYFEFAVQSHGVMDGPFVSVERSTLPDRPSPDLRLS
PLTSTVRLHWCPTPENGEIVEYLILYSNNHTQPEHQWTLITTEGNIPIFAEVHGLSS

misc_feature
1. 23
/gene="Nope"
/notes="unable to sequence region in mRNA"
BASE COUNT 1316 a 1756 c 1634 g 1470 t
ORIGIN
Query Match 100.0%; Score 6176; DB 10; Length 6176;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCGGGGGGACACGGGGCGGGGCTCTGGTGTGTGACCTTCTGCTGCTGCTGCGG 60
Db 1 ATGGCGGGGGGACACGGGGCGGGGCTCTGGTGTGTGACCTTCTGCTGCTGCTGCGG 60
Qy 61 CGGGGGGAGCTGCCATTGCCCCAGGAGACAACCTGTCAAGCTGAGCTGTGATGAGGACCC 120
Db 61 CGGGGGGAGCTGCCATTGCCCCAGGAGACAACCTGTCAAGCTGAGCTGTGATGAGGACCC 120
Qy 121 CTGCAAGTGTATCTGGGCGCTGAGCAGGCTGTGTGTGCTGGACTGCACTTTGGGGGCTACA 180
Db 121 CTGCAAGTGTATCTGGGCGCTGAGCAGGCTGTGTGTGCTGGACTGCACTTTGGGGGCTACA 180
Qy 181 GCTGCTGGGCTCGGACAGGCTGACATGAGCAAGGATGAGACACTGTACTAGAGCAT 240
Db 181 GCTGCTGGGCTCGGACAGGCTGACATGAGCAAGGATGAGACACTGTACTAGAGCAT 240
Qy 241 GAGAACTGTCACCTGTACCCCAATGGCTCTCTGTGGCTGCTCACCCTTAGAGCAAGAA 300
Db 241 GAGAACTGTCACCTGTACCCCAATGGCTCTCTGTGGCTGCTCACCCTTAGAGCAAGAA 300
Qy 301 GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT 360
Db 301 GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT 360
Qy 361 CTGGCCCAAGCGCTAGGAGTGTGGCCAGCCAGGTGCTGTGTTGCTCAAGCTTGCCACA 420
Db 361 CTGGCCCAAGCGCTAGGAGTGTGGCCAGCCAGGTGCTGTGTTGCTCAAGCTTGCCACA 420
Qy 421 CTGGAACACTTCTCTGCAACCCGAGTCCAGAGTCTGGAGGAGAACGGGACAGCAGC 480
Db 421 CTGGAACACTTCTCTGCAACCCGAGTCCAGAGTCTGGAGGAGAACGGGACAGCAGC 480
Qy 481 TTTGAATGCCACACCAAGGGCTTCCAGGCCCTCATCTACTTGGGAAAAGGACAGGTG 540
Db 481 TTTGAATGCCACACCAAGGGCTTCCAGGCCCTCATCTACTTGGGAAAAGGACAGGTG 540
Qy 541 ACCGTGCTGAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCTCCAGATCTTAGAT 600
Db 541 ACCGTGCTGAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCTCCAGATCTTAGAT 600
Qy 601 GTCCAGACAGTGTGATGAGGCTCTCAGCGCTGCGTGGCCCAACCAATTCAGCCCGCAACGA 660
Db 601 GTCCAGACAGTGTGATGAGGCTCTCAGCGCTGCGTGGCCCAACCAATTCAGCCCGCAACGA 660
Qy 661 TTCAAGCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGCTCTTGGAGGGCTTACCAAGGGG 720
Db 661 TTCAAGCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGCTCTTGGAGGGCTTACCAAGGGG 720
Qy 721 CAGGATGTGTCTATTGTTGGCAGCCCAAGAGAACCAACCGGTAGTGTCTGGACAGAAATGTA 780
Db 721 CAGGATGTGTCTATTGTTGGCAGCCCAAGAGAACCAACCGGTAGTGTCTGGACAGAAATGTA 780
Qy 781 GTGATGAGTGTGCTGGCTCTGCTGACCCCAACCCCTTTTGTGTCTCTGGGTCCGACAGAT 840
Db 781 GTGATGAGTGTGCTGGCTCTGCTGACCCCAACCCCTTTTGTGTCTCTGGGTCCGACAGAT 840
Qy 841 GGAAAGCTTATCTCCAGGATGTCTATCGTTCTGGGGCGGACCAATCTACTCTATCGCCAGC 900

[illegible]

1921	ATGGAGTCCCTGGTGTGTGTCATATGGCAGCGCCGCCCTCTACCCACCCAGATCTCTGGATAC	1981
1981	AAACTCTACTGGGAGAGGTGGGAACAAGAGAGGAGGACGATGTGTGACCGCCCCCCAGGG	2041
1981	AAACTCTACTGGGAGAGGTGGGAACAAGAGAGGAGGACGATGTGTGACCGCCCCCCAGGG	2041
2041	GGTCTGTGAGATCAAGCTTTGGAGCGTTCGGGCCCGTTCGGCTTGAAGAGAAAGTGAAGCAG	2101
2041	GGTCTGTGAGATCAAGCTTTGGAGCGTTCGGGCCCGTTCGGCTTGAAGAGAAAGTGAAGCAG	2101
2101	TATGAACATGACCCAGTTAGTCCCTTGGCAGCGCCGTACAGAGTGAAGCTCGTAGCTTTCAAC	2161
2101	TATGAACATGACCCAGTTAGTCCCTTGGCAGCGCCGTACAGAGTGAAGCTCGTAGCTTTCAAC	2161
2161	AAACACGAGGACCGCTACGCTGTGTGTGAAGGGCAGACGAGAGAGCGCCACACGCCA	2221
2161	AAACACGAGGACCGCTACGCTGTGTGTGAAGGGCAGACGAGAGAGCGCCACACGCCA	2221
2221	GACCTGCTATTCAGAGGGGGCCACCGCTGCTCTCTGCCCCATGTCTCACGCAAGTCAAACT	2281
2221	GACCTGCTATTCAGAGGGGGCCACCGCTGCTCTCTGCCCCATGTCTCACGCAAGTCAAACT	2281
2281	AGCTCCACTTCCATTTGGCTTCGGTGGAGAGGACGAGCTTTACCACTGTCAAGATTGTC	2341
2281	AGCTCCACTTCCATTTGGCTTCGGTGGAGAGGACGAGCTTTACCACTGTCAAGATTGTC	2341
2341	AACTACATGTAAGCTTCGGGCCCTCAGGAATGCTTCCCTGGTCAACCTACTAT	2401
2341	AACTACATGTAAGCTTCGGGCCCTCAGGAATGCTTCCCTGGTCAACCTACTAT	2401
2401	ACCAGCTCTGGAGAGACATTTCTCATTTGGCGGCTTGAAACCATTTACCAAGTAGAGTTT	2461
2401	ACCAGCTCTGGAGAGACATTTCTCATTTGGCGGCTTGAAACCATTTACCAAGTAGAGTTT	2461
2461	GGGTACAGTCCACGGAGTGATATGGATGGGCCCTTGGCTCGTGGTAGAGCGTCC	2521
2461	GGGTACAGTCCACGGAGTGATATGGATGGGCCCTTGGCTCGTGGTAGAGCGTCC	2521
2521	ACCTGCCCTGACCGGCCCTTCAACACCTCTTCTGACCTGCGCTGAGCGCCCTGACACCA	2581
2521	ACCTGCCCTGACCGGCCCTTCAACACCTCTTCTGACCTGCGCTGAGCGCCCTGACACCA	2581
2581	TCCACGGTTCCGTTTACACTGGTGTCCCCCAACCGAGCCCAATGGTGGAGATTGTGGAGTAT	2641
2581	TCCACGGTTCCGTTTACACTGGTGTCCCCCAACCGAGCCCAATGGTGGAGATTGTGGAGTAT	2641
2641	CTAATTTCTCAGACACACACACCGAGCCGGAACACAGTGGACACTGCTCACCACA	2701
2641	CTAATTTCTCAGACACACACACCGAGCCGGAACACAGTGGACACTGCTCACCACA	2701
2701	GAGGGAAACATCTTCAGTGCAGAGTCCATGGCTAGAGAGTGACACTCGGTATTCTTC	2761
2701	GAGGGAAACATCTTCAGTGCAGAGTCCATGGCTAGAGAGTGACACTCGGTATTCTTC	2761
2761	AAGATGGGAGCCCGCACAGAGTGGGCCCTTGGCCCTTTTCCCGCTTGCAGATGTGATT	2821
2761	AAGATGGGAGCCCGCACAGAGTGGGCCCTTGGCCCTTTTCCCGCTTGCAGATGTGATT	2821
2821	ACTCTGCAAGAGACATTTCTCAGACTCTTGGATGTGCACCGCGTCAACGCGCATCTCGT	2881
2821	ACTCTGCAAGAGACATTTCTCAGACTCTTGGATGTGCACCGCGTCAACGCGCATCTCGT	2881
2881	GGTGTCTGCTGGGCCCTTCTCTGCTCTCTGGCTGCAATGTGTCTGGCTTACGACAAAGC	2941
2881	GGTGTCTGCTGGGCCCTTCTCTGCTCTCTGGCTGCAATGTGTCTGGCTTACGACAAAGC	2941
2941	TCCCAACGGAAGCCCTTCCCGGATTTGCTCTCTCAGGACACCCACGAGAAACCCAGCGCTC	3001
2941	TCCCAACGGAAGCCCTTCCCGGATTTGCTCTCTCAGGACACCCACGAGAAACCCAGCGCTC	3001
3001	TACACAAGAGTCTCGGCTTGGGCCCTCCAGTGTCCCTGTGCTCCCATGATTTGAGTTCCTC	3061
3001	TACACAAGAGTCTCGGCTTGGGCCCTCCAGTGTCCCTGTGCTCCCATGATTTGAGTTCCTC	3061

Qy	3061	GTGCATCTCTGTCCTCCAGGATTGGTCCCAACCACTTCAGATGTGGAGAACAGCGTGAA	3120
Db	3061	GTGCATCTCTGTCCTCCAGGATTGGTCCCAACCACTTCAGATGTGGAGAACAGCGTGAA	3120
Qy	3121	GTACACAGCCTTATGGGTGGCAGTGTTCAGATTGCCGGGCCCACTCCAGAGAAAGTC	3180
Db	3121	GTACACAGCCTTATGGGTGGCAGTGTTCAGATTGCCGGGCCCACTCCAGAGAAAGTC	3180
Qy	3181	TCCTGGGCTCAGGACAGGGGGACCAAACTGGGCAGGCTCTGGGCAGGCTGTGAGCTGCC	3240
Db	3181	TCCTGGGCTCAGGACAGGGGGACCAAACTGGGCAGGCTCTGGGCAGGCTGTGAGCTGCC	3240
Qy	3241	CAGGTAAGTGTCCAAAGCCCGGCTCTGACCCGTCCTGTCCTCAGCGGGAAACCGGG	3300
Db	3241	CAGGTAAGTGTCCAAAGCCCGGCTCTGACCCGTCCTGTCCTCAGCGGGAAACCGGG	3300
Qy	3301	CAGACACTGCTGCTGCAAGCCCTGGTGTATACCGGCATAAAGAGCAACGGGAGAAAG	3360
Db	3301	CAGACACTGCTGCTGCAAGCCCTGGTGTATACCGGCATAAAGAGCAACGGGAGAAAG	3360
Qy	3361	CCGTCCCAGGCTGCAGGAATCAGGTGGAAGCTCAGGTCACTGTCTCCACTCCGACTTCGGT	3420
Db	3361	CCGTCCCAGGCTGCAGGAATCAGGTGGAAGCTCAGGTCACTGTCTCCACTCCGACTTCGGT	3420
Qy	3421	GCATCCAAAGGATGTCTTGACTCCACTCCAAAGCTCAGGTCACTGTCTCCACTCCGACT	3480
Db	3421	GCATCCAAAGGATGTCTTGACTCCACTCCAAAGCTCAGGTCACTGTCTCCACTCCGACT	3480
Qy	3481	GCAGAGACTCTGCTTCCAGCTCTGGAGCTGTGATCTCAAGAGGAGACTGCGTG	3540
Db	3481	GCAGAGACTCTGCTTCCAGCTCTGGAGCTGTGATCTCAAGAGGAGACTGCGTG	3540
Qy	3541	GGCAGGAGCTGGGAGGCTGCAACCAACCAAGTGGGCGCAGAGAGGCTCACCCTGCTTG	3600
Db	3541	GGCAGGAGCTGGGAGGCTGCAACCAACCAAGTGGGCGCAGAGAGGCTCACCCTGCTTG	3600
Qy	3601	CCAGAAGCAGCAGTGCCTCTGCTCTGCTCAGACTCCAGCCCAAGCAGCTCTATAGAG	3660
Db	3601	CCAGAAGCAGCAGTGCCTCTGCTCTGCTCAGACTCCAGCCCAAGCAGCTCTATAGAG	3660
Qy	3661	GAGGCCCTGGGAAAGCTGCGAGCCCAAGCCCTGTGCTCTTAAAGTCAAGCTCAGCCCAAGC	3720
Db	3661	GAGGCCCTGGGAAAGCTGCGAGCCCAAGCCCTGTGCTCTTAAAGTCAAGCTCAGCCCAAGC	3720
Qy	3721	CTTCCAGGGCCCTGCTCTCTGCTCAGTCTCCAGTGGGCGCAGAGAGGCTCAGCTGCTCA	3780
Db	3721	CTTCCAGGGCCCTGCTCTCTGCTCAGTCTCCAGTGGGCGCAGAGAGGCTCAGCTGCTCA	3780
Qy	3781	GGAACTGCCATGCATGCTTACATGTGTGTACTAGAGATATCCATAAGTCTCTTGGAGC	3840
Db	3781	GGAACTGCCATGCATGCTTACATGTGTGTACTAGAGATATCCATAAGTCTCTTGGAGC	3840
Qy	3841	CTTTAGGGTCTTTGGCTGGGTTGGGAGAACTTTTACTCTCCCTCATATTTCTGCAATCA	3900
Db	3841	CTTTAGGGTCTTTGGCTGGGTTGGGAGAACTTTTACTCTCCCTCATATTTCTGCAATCA	3900
Qy	3901	CATACAGAGGACCTTGAGACACAGCTCTGTGTATGAGACAGCTGTGAAAGTCTGCTGTGT	3960
Db	3901	CATACAGAGGACCTTGAGACACAGCTCTGTGTATGAGACAGCTGTGAAAGTCTGCTGTGT	3960
Qy	3961	GT	4020
Db	3961	GT	4020
Qy	4021	GCCTAGTTGACCCCTCGTGGCAGGATGTGTAAACAGTGTATGAGTGTGTGTGTGTGTGTGT	4080
Db	4021	GCCTAGTTGACCCCTCGTGGCAGGATGTGTAAACAGTGTATGAGTGTGTGTGTGTGTGTGT	4080
Qy	4081	TTTATAGCTTGT	4140
Db	4081	TTTATAGCTTGT	4140

Qy	4141	AAAGACAGCGCCCGCAGCCCTCTTCTGTCTGCTGCTGAGCAGAGTGCCAGTCACTCCAC	4200
Db	4141	AAAGACAGCGCCCGCAGCCCTCTTCTGTCTGCTGCTGAGCAGAGTGCCAGTCACTCCAC	4200
Qy	4201	GGGCTATGACACCGCAGCCTTACACAGCATGCTGTCTATCCCTCGCTTCTTAAAGTTC	4260
Db	4201	GGGCTATGACACCGCAGCCTTACACAGCATGCTGTCTATCCCTCGCTTCTTAAAGTTC	4260
Qy	4261	CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTTTGAGCATCTCTGTACCTAATT	4320
Db	4261	CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTTTGAGCATCTCTGTACCTAATT	4320
Qy	4321	TTGTAATCTGGGAAGTGCCTGGTTGGGAAATCTTCTTTTGCACCCCTGTCTCTCTGTC	4380
Db	4321	TTGTAATCTGGGAAGTGCCTGGTTGGGAAATCTTCTTTTGCACCCCTGTCTCTCTGTC	4380
Qy	4381	CTTCTCTTCAATTTGTTCTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	4440
Db	4381	CTTCTCTTCAATTTGTTCTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	4440
Qy	4441	CTTCTCTTCCCATGATGCCCCCTGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4500
Db	4441	CTTCTCTTCCCATGATGCCCCCTGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4500
Qy	4501	CTTGTCTTTATGCTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4560
Db	4501	CTTGTCTTTATGCTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4560
Qy	4561	ACATGACTACTCTCATGTCT	4620
Db	4561	ACATGACTACTCTCATGTCT	4620
Qy	4621	CTGCAACCCGCTTCTCTGGCAGATGAGCCCACTCTCAAGTAGATCTCTGAAAGAGAGAT	4680
Db	4621	CTGCAACCCGCTTCTCTGGCAGATGAGCCCACTCTCAAGTAGATCTCTGAAAGAGAGAT	4680
Qy	4681	TGGCTTGTCTGCTGCT	4740
Db	4681	TGGCTTGTCTGCTGCT	4740
Qy	4741	GAGTCCAGGCTGCAAGGGGAAACCCCAAGAGGGGTAGGAGTGAAACCAAGAGGCTGA	4800
Db	4741	GAGTCCAGGCTGCAAGGGGAAACCCCAAGAGGGGTAGGAGTGAAACCAAGAGGCTGA	4800
Qy	4801	AAAAAATGCTGCTGCTCACTCTGCAAGAGATGCTGCTCTCTCTCTCTCTCTCTCTCT	4860
Db	4801	AAAAAATGCTGCTGCTCACTCTGCAAGAGATGCTGCTCTCTCTCTCTCTCTCTCTCT	4860
Qy	4861	CTGCTCAAACTGAGGGGTGAGAGAGGGAGCTACTGCGGCTGCTCAATGGCTCAGTTTC	4920
Db	4861	CTGCTCAAACTGAGGGGTGAGAGAGGGAGCTACTGCGGCTGCTCAATGGCTCAGTTTC	4920
Qy	4921	CTGACCTGGAGCCCTGAACTCTGCTTCAAGTAGCAAGAGTCTCTCTCAAGATGCTGT	4980
Db	4921	CTGACCTGGAGCCCTGAACTCTGCTTCAAGTAGCAAGAGTCTCTCTCAAGATGCTGT	4980
Qy	4981	AAGGGAAGTCTTTGCTATAGGAAAGGGCGGCTGCTCAATTTTATTTTCTTTCTTTACA	5040
Db	4981	AAGGGAAGTCTTTGCTATAGGAAAGGGCGGCTGCTCAATTTTATTTTCTTTCTTTACA	5040
Qy	5041	CTGAATCCCAAAATCATCTTACCAAGGGCCCAAGCTGACTGGTATTTCTCTGAGTCAC	5100
Db	5041	CTGAATCCCAAAATCATCTTACCAAGGGCCCAAGCTGACTGGTATTTCTCTGAGTCAC	5100
Qy	5101	AAGAGCCTAGCCATCTCTCTGCTTCTCACTCAGTCTATGCTCCAGAAATGCTCAGATCA	5160
Db	5101	AAGAGCCTAGCCATCTCTCTGCTTCTCACTCAGTCTATGCTCCAGAAATGCTCAGATCA	5160
Qy	5161	GTGCACTCTGCTCTGCTGCTCATCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	5220
Db	5161	GTGCACTCTGCTCTGCTGCTCATCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	5220
Qy	5221	AACCTCTGCTTCTGGGCCAAGTTAGAAATAACCAAGTCTATTTTCTTCTTCTTCTTCTT	5280

Qy	121	CTGCAAGTGTATCTTGGGCGCTTGAGCAGGCTGTGTGTGTGACTGTGACTTTTGGGGCTTACA	180
Db	170	CTGCAAGTGTATCTTGGGCGCTTGAGCAGGCTGTGTGTGTGACTGTGACTTTTGGGGCTTACA	229
Qy	181	GCTCTGGGCTTCCGACAGGGTGACATGGAGCAAGNATGAGACACTGTACTATAGAGAT	240
Db	230	GCTCTGGGCTTCCGACAGGGTGACATGGAGCAAGNATGAGACACTGTACTATAGAGAT	289
Qy	241	GAGAACTGTGACCTGTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAA	300
Db	290	GAGAACTGTGACCTGTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAA	349
Qy	301	GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTTCCTGT	360
Db	350	GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTTCCTGT	409
Qy	361	CTGGCCACAGCCGCTAGGAGTGGTGGCCAGCCAGGTGGTGTGTGTGCTCAAGCTTGGCCACA	420
Db	410	CTGGCCACAGCCGCTAGGAGTGGTGGCCAGCCAGGTGGTGTGTGTGCTCAAGCTTGGCCACA	469
Qy	421	CTCGAAGACTTCTCTCTGACCCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGC	480
Db	470	CTCGAAGACTTCTCTCTGACCCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGC	529
Qy	481	TTTGAATGCCACACCAAGGGCTTCCAGCCCCCATCATTAATCTTGGGAAAAGGACCAAGTG	540
Db	530	TTTGAATGCCACACCAAGGGCTTCCAGCCCCCATCATTAATCTTGGGAAAAGGACCAAGTG	589
Qy	541	ACCGTGCTGAGGAGCCCGGCTCATCATCTTCTCCCAAGTGGCTCTTCCAGATCCTAGAT	600
Db	590	ACCGTGCTGAGGAGCCCGGCTCATCATCTTCTCCCAAGTGGCTCTTCCAGATCCTAGAT	649
Qy	601	GTCAGAGACATGATGCAAGGCTCCTACCGCTGGTGGCCACCAATTCAGCCCGCCCAAGCA	660
Db	650	GTCAGAGACATGATGCAAGGCTCCTACCGCTGGTGGCCACCAATTCAGCCCGCCCAAGCA	709
Qy	661	TTTACGCCAGGAGGCTCTGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACAGAGGG	720
Db	710	TTTACGCCAGGAGGCTCTGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACAGAGGG	769
Qy	721	CAGGATGTGTCTATTTGGGAGCCCCAGAGAACCAACCGTGTGTCTGTGGAAGAAATGTA	780
Db	770	CAGGATGTGTCTATTTGGGAGCCCCAGAGAACCAACCGTGTGTCTGTGGAAGAAATGTA	829
Qy	781	GTGATGAGT	840
Db	830	GTGATGAGT	889
Qy	841	GGAAAGCCTATCTCCAGCGATGTATCGTTCTGGGCGGAGCAATCTACTCATCGCCAGC	900
Db	890	GGAAAGCCTATCTCCAGCGATGTATCGTTCTGGGCGGAGCAATCTACTCATCGCCAGC	949
Qy	901	CGCAGCCTCGGCACTCTGAGTCTATGTCTGCGGAGCCAAACAGCCCTTCAACCGTGC	960
Db	950	CGCAGCCTCGGCACTCTGAGTCTATGTCTGCGGAGCCAAACAGCCCTTCAACCGTGC	1009
Qy	961	TTCCGCACTGCGGCTGTGAGCTCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1020
Db	1010	TTCCGCACTGCGGCTGTGAGCTCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1069
Qy	1021	GAGCGCTCTCGCGAGCGGGCCAGCACCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1080
Db	1070	GAGCGCTCTCGCGAGCGGGCCAGCACCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1129
Qy	1081	CCAGGGCCCGGCTGCACTGGCTGCAAGACGGGATCCCGTTTGGACCAATAGGGCGGCTC	1140
Db	1130	CCAGGGCCCGGCTGCACTGGCTGCAAGACGGGATCCCGTTTGGACCAATAGGGCGGCTC	1189
Qy	1141	AAGGTGACAGGGCGGTGGCGGAGCTTGTGTATCACTCAGATCGGCTGACGAGACGCTGGC	1200
Db	1190	AAGGTGACAGGGCGGTGGCGGAGCTTGTGTATCACTCAGATCGGCTTTCAGGACGCTGGC	1249

Qy	1201	TACTACAGTGTGTAGCAGAAACAGCGCGGGAACTGCTGTGTGCTGCGCTGCGCCCTTGGCG	1260
Db	1250	TACTACAGTGTGTAGCAGAAACAGCGCGGGAACTGCTGTGTGCTGCGCTGCGCCCTTGGCG	1309
Qy	1261	GTAAGTGTGTGCGAGGGGCTGCGAGCGCCCGACTCGGGTCAAGCCACGCGCGCTGAGC	1320
Db	1310	GTAAGTGTGTGCGAGGGGCTGCGAGCGCCCGACTCGGGTCAAGCCACGCGCGCTGAGC	1369
Qy	1321	AGCTCTCTCTGTGTGTGTGCTGGAGCGGCTGTAGTTGACAGCGAGCAAAATCATTTGGC	1380
Db	1370	AGCTCTCTCTGTGTGTGTGCTGGAGCGGCTGTAGTTGACAGCGAGCAAAATCATTTGGC	1429
Qy	1381	TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAATGTGGAGTACCAAGTTTGCAGTA	1440
Db	1430	TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAATGTGGAGTACCAAGTTTGCAGTA	1489
Qy	1441	AACAAATGACACACAGAGCTGCGAGTTCGGGACCTGGAACTTGGAACTTATGAGTTC	1500
Db	1490	AACAAATGACACACAGAGCTGCGAGTTCGGGACCTTGGAACTTGGAACTTATGAGTTC	1549
Qy	1501	TAGTGTGTGCTTACTCTCCAGCTGGGGCCAGCGCAACTCTCAGCCCGAGCCCTTGGTGCAT	1560
Db	1550	TAGTGTGTGCTTACTCTCCAGCTGGGGCCAGCGCAACTCTCAGCCCGAGCCCTTGGTGCAT	1609
Qy	1561	ACACTGGAAGATGTCTCCAGCGCAGCACCCCAAGCTTACTTGTTCAGCCCCCAACCCCTCG	1620
Db	1610	ACACTGGAAGATGTCTCCAGCGCAGCACCCCAAGCTTACTTGTTCAGCCCCCAACCCCTCG	1669
Qy	1621	GACATCAGGCTGCGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1680
Db	1670	GACATCAGGCTGCGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1729
Qy	1681	TACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1740
Db	1730	TACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1789
Qy	1741	AATGAGACACAACTTACCTTAACTCACTTTCAGCCAAACAAAGATGTACCGAGTCCGAT	1800
Db	1790	AATGAGACACAACTTACCTTAACTCACTTTCAGCCAAACAAAGATGTACCGAGTCCGAT	1849
Qy	1801	TCAGCTGCGCATGCTGCTGCTATGAGTGTCTTCTCAGTGTGATGATGATGATGATGATGAT	1860
Db	1850	TCAGCTGCGCATGCTGCTGCTATGAGTGTCTTCTCAGTGTGATGATGATGATGATGATGAT	1909
Qy	1861	GGTGTGCACAAACAGAGCCATGTTTCCCTTTCAGTGTGATGATGATGATGATGATGATGAT	1920
Db	1910	GGTGTGCACAAACAGAGCCATGTTTCCCTTTCAGTGTGATGATGATGATGATGATGATGAT	1969
Qy	1921	ATGAGTGTCTGCTGCTGCTATGAGTGTCTTCTCAGTGTGATGATGATGATGATGATGAT	1980
Db	1970	ATGAGTGTCTGCTGCTGCTATGAGTGTCTTCTCAGTGTGATGATGATGATGATGATGAT	2029
Qy	1981	AAACTCTTACTGCGGAGAGTGGGAAACAGAGAGAGAGAGAGATGATGATGATGATGATGAT	2040
Db	2030	AAACTCTTACTGCGGAGAGTGGGAAACAGAGAGAGAGAGAGATGATGATGATGATGATGAT	2089
Qy	2041	GGTGTGAGATCAAGCTTGGAGCTGCGGCGCTGCGGCTGAGAGAGAGAGAGAGAGAGAG	2100
Db	2090	GGTGTGAGATCAAGCTTGGAGCTGCGGCGCTGCGGCTGAGAGAGAGAGAGAGAGAGAG	2149
Qy	2101	TATGAACTGACCCAGTGTAGTCCCTGCGAGCGCTTACCGCCAGAGTGTGATGATGATGAT	2160
Db	2150	TATGAACTGACCCAGTGTAGTCCCTGCGAGCGCTTACCGCCAGAGTGTGATGATGATGAT	2209
Qy	2161	AAACACAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2220
Db	2210	AAACACAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2269
Qy	2221	GACCTGCTATCCAGAGGGGCGCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2280
Db	2270	GACCTGCTATCCAGAGGGGCGCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2329
Qy	2281	AGCTCTCACTTCCATTTGGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2340

Db	2330	AGCTCCACTTCCATTTTGGCTTCGGTGGAAAGCCAGACCTTTACCACTGTCAAGATTGTC	2389
Qy	2341	AACTACACTGTACGCTTTCGGCCCTCGGGGGCTCAGGAATGCTTCCCTGGTGCACCTACTAT	2400
Db	2390	AACTACACTGTACGCTTTCGGCCCTCGGGGGCTCAGGAATGCTTCCCTGGTGCACCTACTAT	2449
Qy	2401	ACCAGCTCTGGAGAAGACATTTCTCATTTGGCGGCTGAAACCATTTAACAGTACGAGTTT	2460
Db	2450	ACCAGCTCTGGAGAAGACATTTCTCATTTGGCGGCTGAAACCATTTAACAGTACGAGTTT	2509
Qy	2461	GGGGTACAGTCCCAACGAGTGGATATGGATGGGGCCCTTTGGCTCCGCTAGAACGCTCC	2520
Db	2510	GGGGTACAGTCCCAACGAGTGGATATGGATGGGGCCCTTTGGCTCCGCTAGAACGCTCC	2569
Qy	2521	ACCCTGCTGACCGGCCCTTCAACACCTTCCTTGTACCTGCGCTGAGCCGCCCTTGACACCA	2580
Db	2570	ACCCTGCTGACCGGCCCTTCAACACCTTCCTTGTACCTGCGCTGAGCCGCCCTTGACACCA	2629
Qy	2581	TCCAACGTTCCGTTTACACTGGTGTCCCCCAACGAGGCCCAATGGTGGAGATTGTGGAGTAT	2640
Db	2630	TCCAACGTTCCGTTTACACTGGTGTCCCCCAACGAGGCCCAATGGTGGAGATTGTGGAGTAT	2689
Qy	2641	CTAATTCTCTACAGCAACACCAACCCAGCCCGAACAACAGTGGACACTGCTCAACACA	2700
Db	2690	CTAATTCTCTACAGCAACACCAACCCAGCCCGAACAACAGTGGACACTGCTCAACACA	2749
Qy	2701	GAGGGAACATCTTCAGTGCAGAGGTCATATGGCTTAGAGTAGTACACTCGGTATTTCTTC	2760
Db	2750	GAGGGAACATCTTCAGTGCAGAGGTCATATGGCTTAGAGTAGTACACTCGGTATTTCTTC	2809
Qy	2761	AAGATGGGAGCCCGCAACAGAGGTGGGGCCTTGGGCCCTTTTCCGCTTGCAGGATGTGATT	2820
Db	2810	AAGATGGGAGCCCGCAACAGAGGTGGGGCCTTGGGCCCTTTTCCGCTTGCAGGATGTGATT	2869
Qy	2821	ACTCTGAAGAGACATTTCTCAGACTCCTTGGATGTGCACGCGTGCACGGGCATCTCGTG	2880
Db	2870	ACTCTGAAGAGACATTTCTCAGACTCCTTGGATGTGCACGCGTGCACGGGCATCTCGTG	2929
Qy	2881	GGTGTCTGCTGGSCCTTCTCTGCTCCTGGCCCTGACATGTGTGTCGCTACGCAAAAGC	2940
Db	2930	GGTGTCTGCTGGSCCTTCTCTGCTCCTGGCCCTGACATGTGTGTCGCTACGCAAAAGC	2989
Qy	2941	TCCACAGGGAAGCCCTTCCCGGATTTGCTCTCTCAGGCACCCACAGAAACCCAGGCCTC	3000
Db	2990	TCCACAGGGAAGCCCTTCCCGGATTTGCTCTCTCAGGCACCCACAGAAACCCAGGCCTC	3049
Qy	3001	TACACAAGAGCTGGCTTGGCCCTCCAGTGTCCCTGCTGCCCATGATTTGGAGTCCCTC	3060
Db	3050	TACACAAGAGCTGGACTTGGGCCCCCCCAGTGTCCCTGCTGCCCATGATTTGGAGTCCCTC	3109
Qy	3061	GTGCATCTCGTCCCAAGATTTGGTCCCAACCAACCTCAGATGTGGAGCAAGGCTGAA	3120
Db	3110	GTGCATCTCGTCCCAAGATTTGGTCCCAACCAACCTCAGATGTGGAGCAAGGCTGAA	3169
Qy	3121	GTACACAGCCTTATGGGTGGCAGTGTTCAGATTGCGGGGCCACTCCAAAGAAAGATC	3180
Db	3170	GTACACAGCCTTATGGGTGGCAGTGTTCAGATTGCGGGGCCACTCCAAAGAAAGATC	3229
Qy	3181	TCCTGGGCTCAGGCAGGGGGAACAACTGGGCAGGCTCTCTGGGCAGGCTGTGAGCTGCC	3240
Db	3230	TCCTGGGCTCAGGCAGGGGGAACAACTGGGCAGGCTCTCTGGGCAGGCTGTGAGCTGCC	3289
Qy	3241	CAGGGTAGTGGTCCAAGCCGGCTCTGACCGGTGCTGTGCTCTCAGCGGGNACCGGG	3300
Db	3290	CAGGGTAGTGGTCCAAGCCGGCTCTGACCGGTGCTGTGCTCTCAGCGGGNACCGGG	3349
Qy	3301	CAGACACTGTCTGCAAGCCCTGGTGTATGACGGCATAAAGACCAACGGGAGAAAGAG	3360
Db	3350	CAGACACTGTCTGCAAGCCCTGGTGTATGACGGCATAAAGACCAACGGGAGAAAGAG	3409
Qy	3361	CCGTCCTCCAGCTGCAGGAATCAGGTTGGAAAGCTGAGGTCAATTGTCTCACTCCGACTCGGT	3420

Db	3410	CCGTCCCAAGCCCTGCAGGAAATCAGGTGGGAAGCTGAGGTCAATTGTCTCACCCTCGGACTTTCGGT	3469
Qy	3421	GCATCCAAAGGATGTCTCTGACCTCCACCTCCAAAGACCTCGAGCCAGAGGAACCACTGCACT	3480
Db	3470	GCATCCAAAGGATGTCTCTGACCTCCACCTCCAAAGACCTCGAGCCAGAGGAACCACTGCACT	3529
Qy	3481	GCAGAGACTCTGCGCTTCACAGTCTGGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTG	3540
Db	3530	GCAGAGACTCTGCGCTTCACAGTCTGGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTG	3589
Qy	3541	GGCAGGAGCTGGGAGGGTSCCAACAAACAACAGTGGGCGAGAGAGGCTCACTCTGCTTG	3600
Db	3590	GGCAGGAGCTGGGAGGGTSCCAACAAACAACAGTGGGCGAGAGAGGCTCACTCTGCTTG	3649
Qy	3601	CCAGAAGCAGCAGTGCCTCTGCTCTGCTCAGAGCTCCAGGCCAGCAGCTGCTATAGAG	3660
Db	3650	CCAGAAGCAGCAGTGCCTCTGCTCTGCTCAGAGCTCCAGGCCAGCAGCTGCTATAGAG	3709
Qy	3661	GAGGCCCTGGGAAAAGCTGCCAGCCCAAAGCCCTGTGCTCTTAACAGTCAAGCCCAAGC	3720
Db	3710	GAGGCCCTGGGAAAAGCTGCCAGCCCAAAGCCCTGTGCTCTTAACAGTCAAGCCCAAGC	3769
Qy	3721	CTTCCAGGAGCCCTGTCTCTCTGCTCAGGTCCTGAGCCCTGAGCAGAGGAGATATGCTCA	3780
Db	3770	CTTCCAGGAGCCCTGTCTCTCTGCTCAGGTCCTGAGCAGAGGAGATATGCTCA	3829
Qy	3781	GGAACTGCCATGCATGGCTACACATGTGTGTACTAGAGATATCCATAAGTCTCTGGAGC	3840
Db	3830	GGAACTGCCATGCATGGCTACACATGTGTGTACTAGAGATATCCATAAGTCTCTGGAGC	3889
Qy	3841	CTCTTAGGGTCTTTTGGCTGGGGTGGGAGAACTTTTACTCTCCCTCATATTTCTGCATCA	3900
Db	3890	CTCTTAGGGTCTTTTGGCTGGGGTGGGAGAACTTTTACTCTCCCTCATATTTCTGCATCA	3949
Qy	3901	CATACAGGAGGACTTGAGACACAGCTCTGTGTAAATGGACAGGTGTGAGTGTGTGTGT	3960
Db	3950	CATACAGGAGGACTTGAGACACAGCTCTGTGTAAATGGACAGGTGTGAGTGTGTGTGT	4009
Qy	3961	GTGTGTGTGTGTGTCTGCTGGTGGAGCTAGGAAACCTCTCCCTATGTAGCACTCACTGTG	4020
Db	4010	GTGTGTGTGTGTGTCTGCTGGTGGAGCTAGGAAACCTCTCCCTATGTAGCACTCACTGTG	4069
Qy	4021	GCCTAGTTGACCTTCGTGGCAGGATGGTGTAAACAGTGATCAGTGCAGCTCTTTGAGCT	4080
Db	4070	GCCTAGTTGACCTTCGTGGCAGGATGGTGTAAACAGTGATCAGTGCAGCTCTTTGAGCT	4129
Qy	4081	TTTAGCCCTGTGCACCTAGACCTTTTATTAACCTCTGAGAGTGTCTCCAGTGTGTGTCTAC	4140
Db	4130	TTTAGCCCTGTGCACCTAGACCTTTTATTAACCTCTGAGAGTGTCTCCAGTGTGTGTCTAC	4189
Qy	4141	AAAGACAGCCGCCAGCCCTCTCTGTGACGTGTGCTGAGCAGAGTGCAGTCAACTCCAC	4200
Db	4190	AAAGACAGCCGCCAGCCCTCTCTGTGACGTGTGCTGAGCAGAGTGCAGTCAACTCCAC	4249
Qy	4201	GGGCCTATGACACCGCAGCCTACACACAGCATGCTGTCACTCCCTCGGCTCTTAAGTTC	4260
Db	4250	GGGCCTATGACACCGCAGCCTACACACAGCATGCTGTCACTCCCTCGGCTCTTAAGTTC	4309
Qy	4261	CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTTTGAGGATCTCTGTACCTAAAT	4320
Db	4310	CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTTTGAGGATCTCTGTACCTAAAT	4369
Qy	4321	TTGTAAATCTGGGAAGTGCCTGGTTTGGGAAATCTTTCTTTGGCACCTGTCCCTCTGTGC	4380
Db	4370	TTGTAAATCTGGGAAGTGCCTGGTTTGGGAAATCTTTCTTTGGCACCTGTCCCTCTGTGC	4429
Qy	4381	CTTTCTCTTCAATTTGTTCTGGTGTATCTGTCTCATGTGTCAATCTTGTCTGATTAATCTGGGGCC	4440
Db	4430	CTTTCTCTTCAATTTGTTCTGGTGTATCTGTCTCATGTGTCAATCTTGTCTGATTAATCTGGGGCC	4489
Qy	4441	CTTCTCTTCCATGATGCCCTCGAATTCCTCTCATCTGCTGTGTTCATTTCTGTCTGCCATG	4500
Db	4490	CTTCTCTTCCATGATGCCCTCGAATTCCTCTCATCTGCTGTGTTCATTTCTGTCTGCCATG	4549

Qy	4501	CTTGCTTTTATGTGCTGTGTTTCTGTCGTCCTGAGTTCAACCTATGCAACCTTCTCTTAACA	4560
Db	4550	CTTGCTTTTATGTGCTGTGTTTCTGTCCTGAGTTCAACTATGCAACCTTCTCTTAACA	4609
Qy	4561	ACATGACTACTCATGTCTGTTTCAGACACCATAGTGTGACCCCTGGGTCCCAACAGCTCCC	4620
Db	4610	ACATGACTACTCATGTCTGCTTCAGACACCATAGTGTGACCCCTGGGTCCCAACAGCTCCC	4669
Qy	4621	CTGCAACCCGCTTCTCTGGGCAGATGAGCCCATCTCCAAGTAGATCTCGAAAAGACCCCTTG	4680
Db	4670	CTGCAACCCGCTTCTCTGGGCAGATGA-CCCACCTCCAAGTAGATCTCGAAAAGA-CCCTTG	4727
Qy	4681	TGGCTTGTCTGGCTGCGCCTCCCTTGGTGTGTAGATGAGAAGTFTTCTATAGGAAGAGAT	4740
Db	4728	TGGCTTGTCTGGCTGCGCCTCCCTTGGTGTGTAGATGAGAAGGTTTCTATAGGAAGAGAT	4787
Qy	4741	GAGTCCAGGCTGCACAGGGGAACCCCAAGAAAGGGTAGGAGGTGAACCAAGAGGCTGA	4800
Db	4788	GAGTCCAGGCTGCACAGGGGAACCCCAAGAAAGGGTAGGAGGTGAACCAAGAGGCTG-	4846
Qy	4801	AAAAAATGGCTGCACCCCATCTGCACAGAGAGATGGGTGTGTCTTTTGACGTGCAGTC	4860
Db	4847	AAAAAATGGCTGCACCCCATCTGCACAGAGAGATGGGTGTGTCTTTTGACGTGCAGTC	4906
Qy	4861	CTGCTGAAACTGAAGGGGTGAGAGAGGGAGGCTACTGGGGCTGCCATGGCTCAGTTC	4920
Db	4907	CTGCTGAAACTGAAGGGGTGAGAGAGGGAGGCTACTGGGGCTGCCATGGCTCAGTTC	4966
Qy	4921	CTGACCTTGAGCCCTGAACTGCTGCTTCAGAGTAGCAAGAGTTTTCCTCAAAGATGCTGT	4980
Db	4967	CTGACCTTGAGCCCTGAACTGCTGCTTCAGAGTAGCAAGAGTTTTCCTCAAAGATGCTGT	5026
Qy	4981	AAGGGAAGTCTTTGTCATAGAAAAAGGGCGGTGGCTCATTTTATTTTATCTTTCTTACA	5040
Db	5027	AAGGGAAGTCTTTGTCATAGAAAAAGGGCGGTGGCTCATTTTATTTTATCTTTCTTACA	5086
Qy	5041	CTGAATCCCAAAATCATCTTACCAAAAGGGCCAAAGCTGACTGGTATTTCTGTGAGTCAC	5100
Db	5087	CTGAATCCCAAAATCATCTTACCAAAAGGGCCAAAGCTGACTGGTATTTCTGTGAGTCAC	5146
Qy	5101	AAGAGCCATGCACTCTCTGTGTTTCTCACTCTAGTCAATGTCCCAAGAAATGTGAGATCCA	5160
Db	5147	AAGAGCCATGCCATCTCTGTGTTTCTCACTCTAGTCAATGTCCCAAGAAATGTGAGATCCA	5206
Qy	5161	GTGGCATCTGTGCTCTTGCTGCACATCTTTCTATTTCAACTGGCTGGCACATCAAGTGTT	5220
Db	5207	GTGGCATCTGTGCTCTTGCTGCACATCTTTCTATTTCAACTGGCTGGCACATCAAGTGTT	5266
Qy	5221	AACCTGGCTTCTGGCCNAGTTAGAATAACAGAGTCTATTTTCCCTTTATTTTATTTTA	5280
Db	5267	AACCTGGCTTCTGGCCNAGTTAGAATAAACAGAGTCTATTTTCCCTTTATTTTATTTTA	5326
Qy	5281	TTTTTATTTTATTTTATGCTTTTCACTGGAGTTGTAGCTTCTGAAAGGGCTGTGTGTTTATT	5340
Db	5327	TTTTTATTTTATTTTATGCTTTTCACTGGAGTTGTAGCTTCTGAAAGGGCTGTGTGTTTATT	5386
Qy	5341	AGCCTTGTGTGCACTCATGTTTGACCCCAACCCCAATTTCTTCTCTCCCTCTTCAGC	5400
Db	5387	AGCCTTGTGTGCACTCATGTTTGACCCCAACCCCAATTTCTTCTCTCCCTCTTCAGC	5446
Qy	5401	CAGCCTATGATAACACTAAGATTATTTAATGTCTGGCTTCGTACTCATTAAGACAGGAT	5460
Db	5447	CAGCCTATGATAACACTAAGATTATTTAATGTCTGGCTTCGTACTCATTAAGACAGGAT	5506
Qy	5461	TGTCATCTGAACACTACTTCTATAGCATTTCAAAGTGGCCACGGCCAAACACCGATGTTT	5520
Db	5507	TGTCATCTGAACACTACTTCTATAGCATTTCAAAGTGGCCACGGCCAAACACCGATGTTT	5566
Qy	5521	CTTCAITGCTCTGAAGGTCAAGACCTCATTTGTTTTCCTGGTTAGATCTTTTCTCTCC	5580
Db	5567	CTTCAITGCTCTGAAGGTCAAGACCTCATTTGTTTTCCTGGTTAGATCTTTTCTCTCC	5626

Qy	5581	TTGCTTGAATGAATAAACCGTTTAAACAGTAGGCTCTTTAGAGCATCAACACATAGTCAT	5640
Db	5627	TTGCTTGAATGAATAAACCGTTTAAACAGTAGGCTCTTTAGAGCATCAACACATAGTCAT	5686
Qy	5641	TCCTCAGTTCCTGTTTAAACAACGACCTTGAGGTTCTGGGTTTAAATTAATAGCTGCACA	5700
Db	5687	TCCTCAGTTCCTGTTTAAACAACGACCT--GAGGTTCTGGTTTAAATTAATAGCTGCACA	5744
Qy	5701	TGAGACAATTTATAACCCATTAGGCTGGGTGGAAAAATTG--TTCTCAAAAGCAAAATAAGTA	5759
Db	5745	TGAGACAATTTATAACCCATTAGGCTGGGTGGAAAAATTGTTCTCAAAAGCAAAATAAGTA	5804
Qy	5760	ATAAATCTGGTATCTGCTATAAATCACTACAGTTGATTAAGAAAGTAGCAGAACTCACTAGC	5819
Db	5805	ATAAATCTGGTATCTGCTATAAATCACTACAGTTGATTAAGAAAGTAGCAGAACTCACTAGC	5864
Qy	5820	ATTATATATCAATTGGGGTTCCTGAGTAACCTGGGAGAGTGTAGCTTTTGTGACTTTGTAGCAG	5879
Db	5865	ATTATATATCAATTGGGGTTCCTGAGTAACCTGGGAGAGTGTAGCTTTTGTGACTTTGTAGCAG	5924
Qy	5880	CAGGTCCTTAATTAGGAAGAGTCTGTGTGGCTTTTACAGGGCAATTAGTCCCTTTGTGCTTTGC	5939
Db	5925	CAGGTCCTTAATTAGGAAGAGTCTGTGTGGCTTTTACAGGGCAATTAGTCCCTTTGTGCTTTGC	5984
Qy	5940	CATGGATGCTTAAAGTCTTTTGAGTCTCAATTAAGAAATTCCTTTCTGGAAGCATGACA	5999
Db	5985	CATGGATGCTTAAAGTCTTTTGAGTCTCAATTAAGAAATTCCTTTCTGGAAGCATGACA	6044
Qy	6000	AGTGTATCGCAATCACTTACAGTCTCACTCGTTTACCTG--GCTTAGTTTGGCTGGGTTAT	6058
Db	6045	AGTGTAT--GCAATCACTTACAGTCTCACTCGTTTACCTGGTCTAGTTTGTGCTGGGTTAT	6103
Qy	6059	TTAATTGCACCTTCCAGCATCATGCTTCCCTCTTACAAATATCATATCTCTTAATTGTTAC	6118
Db	6104	TTAATTGCACCTTCCAGCATCATGCTTCCCTCTTACAAATATCATATCTCTTAATTGTTAC	6163
Qy	6119	ACTAAGGTGTTGATCATGTATCTGTCCCTGTAAGAAATTAATAAATTTTCCAG	6174
Db	6164	ACTAAGGTGTTGATCATGTATCTGTCCCTGTAAGAAATTAATAAATTTTCCAG	6219
RESULT 4			
AB052621		6222 bp mRNA linear	ROD 14-APR-2001
LOCUS			
DEFINITION	Mus musculus DDM36E mRNA, complete cds.		
ACCESSION	AB052621		
VERSION	AB052621.1	GI:11862940	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (eites)		
TITLE	Murakami,H., Nakamata,T., Nakayama,T., Yamamoto,H., Hosaka,T., Aoyama,T., Nagayama,S., Oka,M., Kiyono,T., Sasaki,M.S., Nakamura,T. and Toguchida,J.		
JOURNAL	Up-regulation of a ras effector and down-regulation of osteoblasts adhesion molecule are associated with transfection of osteoblasts		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 6222)		
TITLE	Toguchida,J., Nakamata,T., Murakami,H., Nakayama,T. and Nakamura,T.		
JOURNAL	Submitted (11-DEC-2000) Junya Toguchida, Kyoto University, Institute for Frontier Medical Sciences; 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto, Kyoto 606-8507, Japan (E-mail:togjun@frontier.kyoto-u.ac.jp, Tel:81-75-751-4134, Fax:81-75-751-4144)		
FEATURES	Location/Qualifiers		
SOURCE	1..6222		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/CBA"		
	/db_xref="taxon:10090"		

[illegible]

QY 5937 TGCCATGATGCTTAAGTTCTTTGAGTCTCAATTAAGAATCTTTTCTCGAAGCATG 5996
 Db 5985 TGCCATGATGCTTAAGTTCTTTGAGTCTCAATTAAGAATCTTTTCTCGAAGCATG 6044
 QY 5997 ACAAGTGTATGCAATACTTACATGCTCACTCGTTTACTG-GCTTAGTTGTGCTGGGT 6055
 Db 6045 ACAAGTGTAT-GCAATATTACATGCTCACTCGTTTACTGCTAGTTGTGCTGGGT 6103
 QY 6056 TATTTAAATGGCACTTCCAGCATCATGCTTCTCTTCAAAATATGATATTCTTTATTGT 6115
 Db 6104 TATTTAAATGGCACTTCCAGCATCATGCTTCTCTTCAAAATATGATATTCTTTATTGT 6163
 QY 6116 TACACTAAGTGTTCATCATGATATCTGCTCCCTGTAAGAAATTAATAAATAATTTCCAG 6174
 Db 6164 TACACTAAGTGTTCATCATGATATCTGCTCCCTGTAAGAAATTAATAAATAATTTCCAG 6222

RESULT 5
 AK122535
 LOCUS AK122535 6301 bp mRNA linear ROD 15-MAR-2003
 DEFINITION Mus musculus mRNA for mKIAA1628 protein.
 ACCESSION AK122535
 VERSION AK122535.1 GI:28972801
 KEYWORDS FLI_CDNA.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S.,
 Nakajima,D., Nagase,T., Ohara,O. and Koga,H.
 Prediction of the coding sequences of mouse homologues of KIAA
 gene: II. The complete nucleotide sequences of 400 mouse
 KIAA-homologous cDNAs identified by screening of terminal sequences
 of cDNA clones randomly sampled from size-fractionated libraries
 DNA Res. 10, 35-48 (2003)
 2 (bases 1 to 6301)
 Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
 Direct Submission
 Submitted (07-FEB-2003) Hiasahi Koga, Kazusa DNA Research
 Institute, Laboratory for Genome Informatics; 2-6-7
 Kazusa-kanatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
 The CREATE program supported by Japan science and technology
 corporation; cDNA full insert sequencing: Kazusa DNA Research
 Institute; cDNA library construction, clone selection and 5'- &
 3'-end one pass sequencing.
 Location/Qualifiers
 1..6301
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="mb:00013"
 /tissue_type="brain"
 /dev_stage="adult"
 /note="vector:modified pBC SK+"
 1..6301
 /gene="mKIAA1628"
 <1..3887
 /gene="mKIAA1628"
 /note="CDS is predicted by in silico analysis. Start codon
 is not identified."
 /codon_start=3
 /evidence=not experimental
 /product="mKIAA1628 protein"
 /protein_id="BAC65817.1"
 /db_xref="GI:28972802"
 /translation="EGSQNPNSAMHGRGELPLPORTTVKLSCDRGLQVILGPQAVV
 LDCTLGATGAPPTVMTSKDQDTVLEHNLHLLPNSLWLSPLEQSDSDDEALRI
 WKVTGTSCLAHPLGVASQVAVKLATLEDFSLHPESQIVENGTFARFCHTKGL
 PAPITKEDQVTPESRLITLPLNGVLOILDQSDAGSYRCVATNSARFQSEAS
 LTVALRGSLEATRGQDVVIAAPENTTVVSGQSVVMBCVASADPTPPFVSVRQDQSK

KELGIQTSGDAALQCGSNLTVGEGVGTVPASPSSLLLLCHQMESLSRMSFPWAG
 PTYSSPARSLGTLIESMAEPTSPARVISPLRLLSSBCLLPQRRRRRRRGRGPAIR
 ASCAGPGLSPRECTCTGSCDPMGRVKGQGGSLVTQIGLADGATYQCVAEHS
 AGTACAAAPLAVVRBGLPBPATRTVATPLSSSVLWAMERPELHSQIIGFSLHYOK
 ARGVDNVYQFVANNDDLEBNTDYFVVAISOLGASRSTSSVLTLDIV
 PSAAPOLJTLSSPNPDIIRVAWLPLPSSLSNGQVLYKYIEYGLKBEQDQVPEVGRNE
 TQTLTNSLQPNKYVRISAGTGAGYQVPSQWQHRTPGVHNSHQVPPAPAEALKYRAK
 MESLIVSWQPPPHPTQISGYKLTVREVGVTBERADGRRPGGRGQDWDVGPVRLKKVY
 KQYELTQLVPRGLYEVKLVAFNKHEGDYAAVMKGTESKAPTDPLOIQRGPPPLPAHVH
 AENSNTSILWLRWKDPFTTKIVNTYTRFGPMGLRNASLVTYTTSSGDIILGGLP
 FTKYEPVAVQSHGMDQPGFVSVRSTLPDRPSTPPSDRLSLVLTPTGTVLHWCPTTB
 PNEIYEVLTLYLNNHTQPEHOMTLTTEGNIISAEVHGLESDTRVPEPKMGATEVGP
 GPSPRLQDVLTQKTFSDSLDVAVTGIIYGVCLGLLCLLACWAGLRSHREALPG
 LSSSGTGNPALYTRAKLGPPSVAHLSLSLHVRPQDMPSPSPSDVEDKAEVHSLMG
 GSVDCRHSKRIKISWAQAGCPNMGWAGCELPQSGSPRALTRALLPAGTQTLTLL
 LQALVYDAIKSNGRKKFSPACRNQVEAEVIVHSDFGASKGCPDHLHLDLPEPPLTAE
 TLESTGAVDLSSQADWLGRGCGOPTTSGPERLTCLPEAASACSCSDLPQSTALE
 EAPKSCOPKALCPLTVPSLPRAPVSSAQPV"

BASE COUNT 1362 a 1770 c 1669 g 1500 t
 ORIGIN

Query Match 94.1%; Score 5810.2; DB 10; Length 6301;
 Best Local Similarity 97.0%; Pred. No. 0;
 Matches 6077; Conservative 0; Mismatches 33; Indels 157; Gaps 9;

QY 62 GCGGGAGCTGCCATTGCCCCAGGAGACAACCTGTCAAGCTGAGCTGTGATGAGGACCCC 121
 Db 40 GAGGGAGCTGCCATTGCCCCAGGAGACAACCTGTCAAGCTGAGCTGTGATGAGGACCCC 99
 QY 122 TGCAGGTGATCTGGGCGCTGAGCAGGCTGTGGTCTGGACTGCACCTTTGGGGCTACAG 181
 Db 100 TGCAGGTGATCTGGGCGCTGAGCAGGCTGTGGTCTGGACTGCACCTTTGGGGCTACAG 159
 QY 182 CTGCTGGGCGCTCCAGCAGAGGTGATGAGCAGAGGTGAGACACTGTACTAGAGCATG 241
 Db 160 CTGCTGGGCGCTCCAGCAGAGGTGATGAGCAGAGGTGAGACACTGTACTAGAGCATG 219
 QY 242 AGAACCTGCACTGCTACCCAAATGGCTCCCTGTGGTGTCTCTACCCCTAGACAGAAG 301
 Db 220 AGAACCTGCACTGCTACCCAAATGGCTCCCTGTGGTGTCTCTACCCCTAGACAGAAG 279
 QY 302 ACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTGTCT 361
 Db 280 ACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTGTCT 339
 QY 362 TGGCCCAAGCCGCTAGGAGTGTGTCAGCAGGCTGTCTGTGGTCAAGCTTCCCAACAC 421
 Db 340 TGGCCCAAGCCGCTAGGAGTGTGTCAGCAGGCTGTCTGTGGTCAAGCTTCCCAACAC 399
 QY 422 TCGAAGACTTCTCTCTGACCCCGAGTCCAGATTTGTGGAGGAGAACCGGACACAGCT 481
 Db 400 TCGAAGACTTCTCTCTGACCCCGAGTCCAGATTTGTGGAGGAGAACCGGACACAGCT 459
 QY 482 TTGAATGCCACACCAAGGGCTTCCAGCCCCCATCATTAATTGGGAAAAGACACAGGTGA 541
 Db 460 TTGAATGCCACACCAAGGGCTTCCAGCCCCCATCATTAATTGGGAAAAGACACAGGTGA 519
 QY 542 CCGTGCCTGAGGAGCCCGGCTCATCACTCTTCCCAAGTGGCTCCTCCAGATCTCTAGATG 601
 Db 520 CCGTGCCTGAGGAGTCCCGGCTCATCACTCTTCCCAAGTGGCTCCTCCAGATCTCTAGATG 579
 QY 602 TCAGGACAGTGTGAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAACGAT 661
 Db 580 TCAGGACAGTGTGAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAACGAT 639
 QY 662 TCAGCCAGGAGGCTCCCTCACTGTGGCCCTCAGAGGGTCTTTTGGAGGCTTACAGGGGGC 721
 Db 640 TCAGCCAGGAGGCTCCCTCACTGTGGCCCTCAGAGGGTCTTTTGGAGGCTTACAGGGGGC 699
 QY 722 AGGATGTGTCATTTGGGAGCCGCCAGACACACAGGTAGTCTCTGGACAGATGTAG 781
 Db 700 AGGATGTGTCATTTGGGAGCCGCCAGACACACAGGTAGTCTCTGGACAGATGTAG 759

Db	2920	GGCCCTTTTCCCGCTTGCGAGATGTGATTA	CTGTGCAAAAAGACATTTCTCAGACTCTCTGG	2979
Qy	2852	ATGTGCACGCGCTCACGGGATCATCTGTGGTGTCTGCCTGGGCTTCTCTGCTCTCTGG	2911	
Db	2980	ATGTGCACGCGCTCACGGGATCATCTGTGGTGTCTGCCTGGGCTTCTCTGCTCTCTGG	3039	
Qy	2912	CTTGCAATGTGTCTGTGGCTTACGACAAAGCTCCCA	CAGGGAAGCCCTTCCCGGAAATGTGTCT	2971
Db	3040	CCTGCATGTGTGTCTGGCTTACGACGAAGCTCCCA	CAGGGAAGCCCTTCCCGGAAATGTGTCT	3099
Qy	2972	CCTCAGGACCCCAAGGAAACCCACGCTCTTACACA	CAAGAGCTCCGCTTGGGCTCCCAAGT	3031
Db	3100	CCTCAGGACCCCAAGGAAACCCACGCTCTTACACA	CAAGAGCTCCGCTTGGGCTCCCAAGT	3159
Qy	3032	TCCCTGTCTGCCATGAGTTGGAGTCCCTCGTGATCTCT	CGTCCCAAGATTTGGTCCCCAC	3091
Db	3160	TCCCTGTCTGCCATGAGTTGGAGTCCCTCGTGATCTCT	CGTCCCAAGATTTGGTCCCCAC	3219
Qy	3092	CACCTCAGATGTGGAAAGACAAGGCTGAAGTACA	CAGCCTTATGGGTGGCAGTGTGTTCAG	3151
Db	3220	CACCTCAGATGTGGAAAGACAAGGCTGAAGTACA	CAGCCTTATGGGTGGCAGTGTGTTCAG	3279
Qy	3152	ATTGCCGGGCCCCTCCAAAGAAAGATCTCTGGGCT	CAGGCAAGGGGACCAAACTGGG	3211
Db	3280	ATTGCCGGGCCCCTCCAAAGAAAGATCTCTGGGCT	CAGGCAAGGGGACCAAACTGGG	3339
Qy	3212	CAGGCTCCTGGGCAAGGCTGTGAGCTGCCCAGGGT	TAGTGTGTCAAAGCCCGGCTCTGACCC	3271
Db	3340	CAGGCTCCTGGGCAAGGCTGTGAGCTGCCCAGGGT	TAGTGTGTCAAAGCCCGGCTCTGACCC	3399
Qy	3272	GTGCTCTGTCTGCCCTCCAGCGGGAACCGGGCAG	ACACTGTCTGTCTGCAAGCCCTGGTGTATG	3331
Db	3400	GTGCTCTGTCTGCCCTCCAGCGGGAACCGGGCAG	ACACTGTCTGTCTGCAAGCCCTGGTGTATG	3459
Qy	3332	ACGGCATAAAGAGCAACCGGAGAAAGACCGCTCC	CAGCCTGCAAGGAAATCAGGTGGAAG	3391
Db	3460	ACGGCATAAAGAGCAACCGGAGAAAGACCGCTCC	CAGCCTGCAAGGAAATCAGGTGGAAG	3519
Qy	3392	CTGAGTCAATTGTCCACTCCGACTTCGAGTGTGAT	CTCAAAAGATGTCTGACTCCACCTCC	3451
Db	3520	CTGAGTCAATTGTCCACTCCGACTTCGAGTGTGAT	CTCAAAAGATGTCTGACTCCACCTCC	3579
Qy	3452	AAGACCTTGGAGCCAGAGGAAACCACTGACTGC	AGAGACTCTGCCCTTCAAGCTTGGAGCTG	3511
Db	3580	AAGACCTTGGAGCCAGAGGAAACCACTGACTGC	AGAGACTCTGCCCTTCAAGCTTGGAGCTG	3639
Qy	3512	TGGATCTGTCTCAAGGAGGAGACTGGCTGGG	CAGGAGCTGGGAGGCTGCCAACAA	3571
Db	3640	TGGATCTGTCTCAAGGAGGAGACTGGCTGGG	CAGGAGCTGGGAGGCTGCCAACAA	3699
Qy	3572	CCAGTGGGCGAGAGAGGCTCACCTGTGCGCAG	AAGCAGCAGCTGCTCCCTGCTCTGCT	3631
Db	3700	CCAGTGGGCGAGAGAGGCTCACCTGTGCGCAG	AAGCAGCAGCTGCTCCCTGCTCTGCT	3759
Qy	3632	CAGACCTCCAGCCAGCACTGCTATAGAGGAG	CGCCCTTGGGAAAGCTGGCAGCCAAAG	3691
Db	3760	CAGACCTCCAGCCAGCACTGCTATAGAGGAG	CGCCCTTGGGAAAGCTGGCAGCCAAAG	3819
Qy	3692	CCCTGTGTCTTAAAGTACAGCCCTTCCAGGG	CCCTGTCTCTCTGCTCAGG	3751
Db	3820	CCCTGTGTCTTAAAGTACAGCCCTTCCAGGG	CCCTGTCTCTCTGCTCAGG	3879
Qy	3752	TCCCTCAGCAGAGGAGATATGGCTCAGGAA	CATGCCATGCCATGCCATACATGTGTG	3811
Db	3880	TCCCTCAGCAGAGGAGATATGGCTCAGGAA	CATGCCATGCCATGCCATACATGTGTG	3939
Qy	3812	TACTAGAGATATCAATAAGTCTTGGAGCCTCTT	TAGGCTCTTGGGCTGGGCTGGGAG	3871
Db	3940	TACTAGAGATATCAATAAGTCTTGGAGCCTCTT	TAGGCTCTTGGGCTGGGCTGGGAG	3999
Qy	3872	AACTTTACTCTCCCTCATATTCTGCATCA	CATACACAGAGGAGCTTGAGACACAGCTCTGT	3931
Db	4000	AACTTTACTCTCCCTCATATTCTGCATCA	CATACAGAGGAGCTTGAGACACAGCTCTGT	4059

Qy	5012	TGCGTCATTTTATTTTTATCTTTTCTTTACACTGAATCCCAAAATCATCTTACACAAGGG	5071
Db	5138	TGGGTCATTTTATTTTTATCTTTTCTTTACACTGAATCCCAAAATCATCTTACACAAGGG	5197
Qy	5072	CCAAGCCTGACTGGTATTCTTGANGTCAAGAAGGCATGCCATCTCTCTGTGTTCTCACCC	5131
Db	5198	CCAAGCCTGACTGGTATTCTTGAGTCAACAAGAGCCATGCGCATCTCTGTGGTTTCTCACCC	5257
Qy	5132	TCAGTCAATGCCAGAAATGTGCAGATCCAGTGGCATCTGTGCTCTTGCTGCACATCTTTC	5191
Db	5258	TCAGTCAATGCCAGAAATGTGCAGATCCAGTGGCATCTGTGCTCTTGCTGCACATCTTTC	5317
Qy	5192	TATTTCAACTGGGTGGGCATCAAGTGTTHAATCTCTGGCTTCTGGGCGAAGTTAGAAATAA	5251
Db	5318	TATTTCAACTGGGTGGGCATCAAGTGTTHAATCTCTGGCTTCTGGGCGAAGTTAGAAATAA	5377
Qy	5252	CCAGTCTATTTTCCCCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	5311
Db	5378	CCAGTCTATTTTCCCCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	5437
Qy	5312	TGTAGCTTCTGAAAGCGTCTGTGTTTATTAGCCCTGTGTGTCACCTCATGTTTGACCCCCAC	5371
Db	5438	TGTAGCTTCTGAAAGCGTCTGTGTTTATTAGCCCTGTGTGTCACCTCATGTTTGACCCCCAC	5497
Qy	5372	CCACATTTCCCTTCTCTCCCTCTTTCAGCCAGCCTATGATAACACATAAAGATTTTAATG	5431
Db	5498	CCACATTTCCCTTCTCTCCCTCTTTCAGCCAGCCTATGATAACACATAAAGATTTTAATG	5557
Qy	5432	CTGGCTTCGTATCTCATTTAAAGACAGGATTTGTCACTTGAACACTCTTCTATAGCATTTCAA	5491
Db	5558	CTGGCTTCGTATCTCATTTAAAGACAGGATTTGTCACTTGAACACTCTTCTATAGCATTTCAA	5617
Qy	5492	GTTGCCACGCGCAACACCACCGTATGTTCTTCATTGCTCTGAAAGGTCAAGAGCCTCATTT	5551
Db	5618	GTTGCCACGCGCAACACCACCGTATGTTCTTCATTGCTCTGAAAGGTCAAGAGCCTCATTT	5677
Qy	5552	TTGTTTCTCTGGTTTAGATTCCTTTCTCTCTGCTTGAATGAAATAACCGTTTTAAACAGT	5611
Db	5678	TTGTTTCTCTGGTTTAGATTCCTTTCTCTCTGCTTGAATGAAATAACCGTTTTAAACAGT	5737
Qy	5612	AGGCTCTTAGCATCACACATAGTCATTTCTCATGTCTCTGTTTAAACAGCACCTTGAG	5671
Db	5738	AGGCTCTTAGCATCACACATAGTCATTTCTCATGTCTCTGTTTAAACAGCACCT - GA	5795
Qy	5672	GTTCTGGGTTTAAATTAATAGCTGCAAAATGAGACAAATTTATAACCCATTTAGGCTGGGTG	5731
Db	5796	GTTCTGGGTTTAAATTAATAGCTGCAAAATGAGACAAATTTATAACCCATTTAGGCTGGGTG	5855
Qy	5732	GAAAATTG-TTCTCAAAAGCAAAATAGTAATAAATCTGSTATCTGCTATAAATCAACAGT	5790
Db	5856	GAAAATTGTTTCTCAAAAGCAAAATAGTAATAAATCTGSTATCTGCTATAAATCAACAGT	5915
Qy	5791	TGATAAGAAAGTAGCCGAACTCACTAGCATTTATATGATTTGGGTTTCTGAGTAACCTGG	5850
Db	5916	TGATAAGAAAGTAGCCGAACTCACTAGCATTTATATGATTTGGGTTTCTGAGTAACCTGG	5975
Qy	5851	GGAGTGTAGCTTTGTGACCTTTGPAGCACACAGGTCTTTATTAGGAAAGTCTGTTGGCTTT	5910
Db	5976	GGAGTGTAGCTTTGTGACCTTTGPAGCACACAGGTCTTTATTAGGAAAGTCTGTTGGCTTT	6035
Qy	5911	TACAGGGCATTAGTCCCTTTGTGCTTTGCCATGGATGCCCTTAAGTTCTTTTGGAGTCTCAT	5970
Db	6036	TACAGGGCATTAGTCCCTTTGTGCTTTGCCATGGATGCCCTTAAGTTCTTTTGGAGTCTCAT	6095
Qy	5971	TTAAGAAATTCCTTTTCTGGAAGCATGACAAGTGTATCGCAATFACTTACATGCTCATCGT	6030
Db	6096	TTAAGAAATTCCTTTTCTGGAAGCATGACAAGTGTAT - GCAATACTTACATGCTCATCGT	6154
Qy	6031	TTACCTG-GCTTAGTGTGCTGGTATTATTAATTTGCACCTTCCAGCATCATGCTTCCCTC	6089
Db	6155	TTACCTGTGCTTAGTGTGCTGGTATTATTAATTTGCACCTTCCAGCATCATGCTTCCCTC	6214
Qy	6090	CTTACAAATATGATATTCTTTATTGTGTACACTAAGGTGTTGCATCATGTATCTGTCCCTGT	6149

Db	6215	CTTACAAATATGATAATTCTTTATTGTGTACACTAAGGGTGTTGATCAATGATCTGTCCCTGT	6274
Qy	6150	AAAGAATTAAATAAACTATTTTCAGAC	6176
Db	6275	AAAGATTAATAAACTATTTTCAGAC	6301
RESULT 6			
AX191303			
LOCUS	AX191303	3756 bp	DNA linear PAT 15-AUG-2001
DEFINITION	Sequence 45 from Patent WO0149714.		
ACCESSION	AX191303		
VERSION	AX191303.1	GI:15209554	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;		
TITLE	1 Salbaum, J. M.		
JOURNAL	Nope polypeptides, encoding nucleic acids and methods of use		
FEATURES	Patent: WO 0149714-A 45 12-JUL-2001; Neurosciences Research Foundation Inc. (US) Location/Qualifiers 1..3756 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090"		
BASE COUNT	774 a 1167 c 1105 g 710 t		
ORIGIN			
Query Match	60.8%; Score 3756; DB 6; Length 3756;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 3756; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	ATGGCGCGGCGGACACAGGCGCGGGCTCTGTGTCTGACCTTCTGCCTGTCTGTCTGCCG	60
Db	1	ATGGCGCGGCGGACACAGGCGCGGGCTCTGTGTCTGACCTTCTGCCTGTCTGTCTGCCG	60
Qy	61	CGCGGGGAGCTGCCATTGTGCCCGACGACAACTGTCAAGCTGAGCTGTGATGAGGGACCC	120
Db	61	CGCGGGGAGCTGCCATTGTGCCCGACGACAACTGTCAAGCTGAGCTGTGATGAGGGACCC	120
Qy	121	CTGCAAGTGATCTCTGGGCGCTGACAGGCTGTGTGTCTGGA CTGCAC TTTGGGGGCTACA	180
Db	121	CTGCAAGTGATCTCTGGGCGCTGACAGGCTGTGTGTCTGGA CTGCAC TTTGGGGGCTACA	180
Qy	181	GCTGTGGGCTCCGACAGGGTGACATGGAGCAAGGATGAGACACTGTACTAGAGCAT	240
Db	181	GCTGTGGGCTCCGACAGGGTGACATGGAGCAAGGATGAGACACTGTACTAGAGCAT	240
Qy	241	GAGAACCTGCACCTGTACC AATGGCTTCCCTGTGGCTGTCTCACCCCTAGACCAAGAA	300
Db	241	GAGAACCTGCACCTGTACC AATGGCTTCCCTGTGGCTGTCTCACCCCTAGACCAAGAA	300
Qy	301	GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT	360
Db	301	GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT	360
Qy	361	CTGGCCCCACAGCCGCTAGGAGTGGTGGCCAGCAGGTGTGTGTGCTCAAGCTTGCCACA	420
Db	361	CTGGCCCCACAGCCGCTAGGAGTGGTGGCCAGCAGGTGTGTGTGCTCAAGCTTGCCACA	420
Qy	421	CTCGAAGACTTCTCTCTGCACCCGAGTCCCAGATTGTGGAGGAGAACGGGACAGCAGC	480
Db	421	CTCGAAGACTTCTCTCTGCACCCGAGTCCCAGATTGTGGAGGAGAACGGGACAGCAGC	480
Qy	481	TTTGAA TGCCACACCAAGGGGCTTCCAGCCCCCATTCATTACTTGGGAAAAGGACAGGTG	540
Db	481	TTTGAA TGCCACACCAAGGGGCTTCCAGCCCCCATTCATTACTTGGGAAAAGGACAGGTG	540
Qy	541	ACCGTGCTGAGGAGCCCCGGCTCATCACTCTTCCCCAAGTGGCTCTCTCCAGATCCTAGAT	600

Db 541 ACCGTGCTGAGGAGCCCGGCTCATCACTCTTCCCAAGTGCTCTCCAGATCCTAGAT 600
Qy 601 GTCCAGACAGTGTATGACGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCAAACGA 660
Db 601 GTCCAGACAGTGTATGACGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCAAACGA 660
Qy 661 TTCAAGCCAGAGGCTCTGCTCACTGTGGCCCTCAGAGGCTCTTTGGAGGCTTACCAAGGGG 720
Db 661 TTCAAGCCAGAGGCTCTGCTCACTGTGGCCCTCAGAGGCTCTTTGGAGGCTTACCAAGGGG 720
Qy 721 CAGATGTGGTCACTTGTGGAGCCCGCCAGAGAACCAACGCTAGTGTCTGGACAGATGTA 780
Db 721 CAGATGTGGTCACTTGTGGAGCCCGCCAGAGAACCAACGCTAGTGTCTGGACAGATGTA 780
Qy 781 GTGATGGAGTGCTGGCTCTGCTGACCCCAACCCCTTTTGTGTCTTGGTCCGACAGGAT 840
Db 781 GTGATGGAGTGCTGGCTCTGCTGACCCCAACCCCTTTTGTGTCTTGGTCCGACAGGAT 840
Qy 841 GGAAAGCCTATCTCCACGATGTCAATGCTTCTGGGCGGACCAATCTACTCATCGCCAGC 900
Db 841 GGAAAGCCTATCTCCACGATGTCAATGCTTCTGGGCGGACCAATCTACTCATCGCCAGC 900
Qy 901 GCGCAGCCTCGGCACTCTGAGTCTATGTCTGCGAGCCACAGCCCTCACCGGTGAC 960
Db 901 GCGCAGCCTCGGCACTCTGAGTCTATGTCTGCGAGCCACAGCCCTCACCGGTGAC 960
Qy 961 TTCCGCCACTCGGCTGCTGAGCTCCGAGTGTCTGCGCCCGACGCACTCTCGCAGGCAACC 1020
Db 961 TTCCGCCACTCGGCTGCTGAGCTCCGAGTGTCTGCGCCCGACGCACTCTCGCAGGCAACC 1020
Qy 1021 GAGCGCTCTCGCGGACGCGGCGCAGCACCGCGGCTCTGCTGTCGCGGCGTCCGCGGAG 1080
Db 1021 GAGCGCTCTCGCGGACGCGGCGCAGCACCGCGGCTCTGCTGTCGCGGCGTCCGCGGAG 1080
Qy 1081 CCAGCGCCCGGCTGCTGAGTCTGAGTCTGCGAGTGTCTGCGCCCGACGCACTCTCGCAGGCAACC 1140
Db 1081 CCAGCGCCCGGCTGCTGAGTCTGAGTCTGCGAGTGTCTGCGCCCGACGCACTCTCGCAGGCAACC 1140
Qy 1141 AAGGTGACGAGCGCTGGCGGAGCTTGGTCACTCACTCAGATCGGCTGCGAGGAGCTGGC 1200
Db 1141 AAGGTGACGAGCGCTGGCGGAGCTTGGTCACTCACTCAGATCGGCTGCGAGGAGCTGGC 1200
Qy 1201 TACTACCAAGCTGTAGCAAGAAACAGCGCGGAACTGCTGTGCGCTGCGCCCTGGCG 1260
Db 1201 TACTACCAAGCTGTAGCAAGAAACAGCGCGGAACTGCTGTGCGCTGCGCCCTGGCG 1260
Qy 1261 GTAGTGTGCGGAGGCGCTGCCAGCGCCCGAGCTGGGTCAAGCCACGCGCTGAGC 1320
Db 1261 GTAGTGTGCGGAGGCGCTGCCAGCGCCCGAGCTGGGTCAAGCCACGCGCTGAGC 1320
Qy 1321 AGCTCCTGTGTGGTGGCTGGAGCGGCTGAGTTGACAGCGAGCAATCAATTGGC 1380
Db 1321 AGCTCCTGTGTGGTGGCTGGAGCGGCTGAGTTGACAGCGAGCAATCAATTGGC 1380
Qy 1381 TTCTCTCTTCACTACAAAGGCAAGGAGTGGACATGTGGAGTACAGTTTGGAGTA 1440
Db 1381 TTCTCTCTTCACTACAAAGGCAAGGAGTGGACATGTGGAGTACAGTTTGGAGTA 1440
Qy 1441 AACATGACACACAGAGTGTGAGTTGGAGCTTGGAACTGGAAACCAACAGGATATGATTC 1500
Db 1441 AACATGACACACAGAGTGTGAGTTGGAGCTTGGAACTGGAAACCAACAGGATATGATTC 1500
Qy 1501 TACGTGTGGCTACTCCAGCTGGGGCCAGCGAACCTCCAGCCAGCCCTGGTGCAT 1560
Db 1501 TACGTGTGGCTACTCCAGCTGGGGCCAGCGAACCTCCAGCCAGCCCTGGTGCAT 1560
Qy 1561 ACATGGACGATGTCCCGAGCGGACGACCCAGCTTACTTGTCTCAGGCCCAACCCCTCG 1620
Db 1561 ACATGGACGATGTCCCGAGCGGACGACCCAGCTTACTTGTCTCAGGCCCAACCCCTCG 1620
Qy 1621 GACATCAGGGTGGAATGGCTGCCCTGCCCTCCAGCTGAGCAATGGACAGGTCTGAAG 1680

Db 1621 GACATCAGGGTGGATGGCTGCCCTCGCTCCAGCTGAGCAATGGACAGGTGCTGAAG 1680
Qy 1681 TACAAGATACAGTACCGTTTGGGGAAGGAAGATCAGGTTTCTCCACCGAGGTGCTGGA 1740
Db 1681 TACAAGATACAGTACCGTTTGGGGAAGGAAGATCAGGTTTCTCCACCGAGGTGCTGGA 1740
Qy 1741 AATGAGACCAACTTACGTTAAACTCACTTCAGCCAAACAAGTGTACCGAGTCCGAT 1800
Db 1741 AATGAGACCAACTTACGTTAAACTCACTTCAGCCAAACAAGTGTACCGAGTCCGAT 1800
Qy 1801 TCAGCTGGCACTGGCGCTGCTATGAGTCCCTTCTCAGTGGATGACAGCAAGACACCT 1860
Db 1801 TCAGCTGGCACTGGCGCTGCTATGAGTCCCTTCTCAGTGGATGACAGCAAGACACCT 1860
Qy 1861 GGTGTGCAAAACAGAGCCATGTTCTTGGCCCTCAGAAATGAAAGGTGAGGCGAAAG 1920
Db 1861 GGTGTGCAAAACAGAGCCATGTTCTTGGCCCTCAGAAATGAAAGGTGAGGCGAAAG 1920
Qy 1921 ATGGAGTCCCTGCTGCTGCTATGGCAGCCGCCCCCTCACCCACCCAGATCTCTGGATAC 1980
Db 1921 ATGGAGTCCCTGCTGCTGCTATGGCAGCCGCCCCCTCACCCACCCAGATCTCTGGATAC 1980
Qy 1981 AAACCTTACTGGGAGAGGTGGAAACAGAGGAGGAGCAGATGCTGACCGCCCCCAGGG 2040
Db 1981 AAACCTTACTGGGAGAGGTGGAAACAGAGGAGGAGCAGATGCTGACCGCCCCCAGGG 2040
Qy 2041 GGTGTGAGATCAAGCTTGGGAGCTCGGCGCTGCGGCTGGAAGAAAGTGAAGCAG 2100
Db 2041 GGTGTGAGATCAAGCTTGGGAGCTCGGCGCTGCGGCTGGAAGAAAGTGAAGCAG 2100
Qy 2101 TATGAACTGACCCAGTTAGTCCCTGGCAGCGCTGACAGGTGAAGCTCTGAGCTTTCAAC 2160
Db 2101 TATGAACTGACCCAGTTAGTCCCTGGCAGCGCTGACAGGTGAAGCTCTGAGCTTTCAAC 2160
Qy 2161 AAACAGAGAGCGCTAGCTGCTGTGGAAGGAGGAGGAGCGAGAGGCGCCACCGCA 2220
Db 2161 AAACAGAGAGCGCTAGCTGCTGTGGAAGGAGGAGGAGCGAGAGGCGCCACCGCA 2220
Qy 2221 GACCTGCTATCCAGAGGGGGCCACCGCTGCTCTGCTGCTCCAGCAGAGTCAAAAC 2280
Db 2221 GACCTGCTATCCAGAGGGGGCCACCGCTGCTCTGCTGCTCCAGCAGAGTCAAAAC 2280
Qy 2281 AGCTCCACTTCCATTTGGCTTGGTGAAGAGCAGACTTTTACACTGTCAAGATTGTC 2340
Db 2281 AGCTCCACTTCCATTTGGCTTGGTGAAGAGCAGACTTTTACACTGTCAAGATTGTC 2340
Qy 2341 AACTACACTGTAGCTTGGCGCCCTGGGGGCTCAGGAAATGCTTCCCTGGTCACTTAT 2400
Db 2341 AACTACACTGTAGCTTGGCGCCCTGGGGGCTCAGGAAATGCTTCCCTGGTCACTTAT 2400
Qy 2401 ACCAGCTCTGGAGAAAGACATTTCTATTTGGGGGCTGAAACCAATTTACCAAGTACGAGTTT 2460
Db 2401 ACCAGCTCTGGAGAAAGACATTTCTATTTGGGGGCTGAAACCAATTTACCAAGTACGAGTTT 2460
Qy 2461 GCGGTACAGTCCACGAGTGGATATGGATGGGCGCTTTGGCTCCGTGTAAGACGTTC 2520
Db 2461 GCGGTACAGTCCACGAGTGGATATGGATGGGCGCTTTGGCTCCGTGTAAGACGTTC 2520
Qy 2521 ACCCTGCTGACCGGCTTCAACACCTCTTCTGACCTGGCCTGAGCGCCCTGACACCA 2580
Db 2521 ACCCTGCTGACCGGCTTCAACACCTCTTCTGACCTGGCCTGAGCGCCCTGACACCA 2580
Qy 2581 TCCACCTTGGTGTACACTGCTGTGCTCCCAAGAGCCCAATGTGAGATTTGTGAGTAT 2640
Db 2581 TCCACCTTGGTGTACACTGCTGTGCTCCCAAGAGCCCAATGTGAGATTTGTGAGTAT 2640
Qy 2641 CTAATCTTACAGCAACCAACACCGCCGAAACCAAGTGGACACTGCTCACCA 2700
Db 2641 CTAATCTTACAGCAACCAACACCGCCGAAACCAAGTGGACACTGCTCACCA 2700
Qy 2701 GAGGGAACATCTTCACTGAGAGGTCCATGGCTTCAAGAGTGAACACTCGTATTTCTTC 2760
Db 2701 GAGGGAACATCTTCACTGAGAGGTCCATGGCTTCAAGAGTGAACACTCGTATTTCTTC 2760

TITLE	Nope polypeptides, encoding nucleic acids and methods of use			
	Patent: WO 0149714-A 3 12-JUL-2001;			
	Neurosciences Research Foundation Inc. (US)			
FEATURES	Location/Qualifiers			
	1..2796			
	/organism="Mus musculus"			
CDS	/mol_type="genomic DNA"			
	/db_xref="taxon:10090"			
	<1..>2796			
	/notes="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAC51121.1"			
	/translation="GSLPQETTVKLSCEGLPQVILGPQAVLVLDCTLGATAAPP			
	TAVTKGQDVLVLEHNLHLLPNSLWLSPLEQSDDEBALRIWKTEGYSCLAH			
	SFLGVASQAVVVKLALDLSLHPSQIVENGRTARFCHTKGLPAPITWKEQVTR			
	VPEEPLRLITLQILDQSDAGSYECVATNSARQSFQSEALTVALLRGSBATR			
	GDQVITVAAPENTVYVQNVMECVASDPTFPVSWRQDKPISTDIVILGRNLL			
	IASAQRHGGVTVCRANKPLTRDFATAAELRVLAAPASQAPALSRTRASTAFVC			
	RASGEPRPALHWDGILPLRPNRGRVQGGSLVITQIGLODAGYQCVASNSAGTA			
	CAAPLAVVVRGLPSAPTRVATPLSSSLVAVMRBELHSEQLIGFSLHYQKRGV			
	DNVEYFVNDTTELQVRDLAPNTDYEFVVAISQIGASRTSSPALVHTLDDVSA			
	PQLTSSPNPSDIRAWLPLPSLSNGQVLYKIEYGLKEDQVSTFVPGNETQLTL			
	NSLQNPVRVIRISAGTGAGYGVPSQWMOHRTKTPGVHNSHPVFAELKVRAMESLV			
	VSNQPPHPPTQISGYKLYMEVDGEAEDRPPGRGQDQWVGVRLLKKVQVYEL			
	TOLVGRPYEVKLVAFNKEDGVAAVWKTEKAPTDLPIORGPEPLPAHYHARSNS			
	STSLNLRWKKPDETTVKIYNYTRFGPWGLRNALVITYTSSGREDILIGLAPPTKYE			
	FAVQSHGMDGPFQSVSRSTLPPDRPSPSLRSLPTSPTRVHMCPTPEPNGEI			
	VEYLIISSNNHTOPEHOMTLTTEGNIIFSAEVLHLESDFRYFFKMGARTFVGGPFSR			
	LDVITLQETFSLSLDVH"			
BASE COUNT		591 a	849 c	821 g 535 t
ORIGIN				
Query Match		45.3%; Score 2796; DB 6; Length 2796;		
Best Local Similarity		100.0%; Pred. No. 0;		
Matches 2796; Conservative		0; Mismatches 0; Indels 0; Gaps 0;		
QY	64	GGGGAGCTGCTGCTGGCCCGCAGGAGACAACTGTCAAGCTGAGCTGTCATGAGGAGCCCTG	123	
DB	1	GGGGAGCTGCTGCTGGCCCGCAGGAGACAACTGTCAAGCTGAGCTGTCATGAGGAGCCCTG	60	
QY	124	CAAGTGATCTGGGCGCTGAGCAGGCTGTGGTCTGGAGCTGCACTTTGGGGGCTACAGCT	183	
DB	61	CAAGTGATCTGGGCGCTGAGCAGGCTGTGGTCTGGAGCTGCACTTTGGGGGCTACAGCT	120	
QY	184	GCTGGGCTCCGACGAGGCTGACATGAGAGCAAGATGGAGACACTGTATAGAGATGAG	243	
DB	121	GCTGGGCTCCGACGAGGCTGACATGAGAGCAAGATGGAGACACTGTATAGAGATGAG	180	
QY	244	AACCTGCACCTGTACCAATGCTCCCTGTGGCTGTCTCACCCTAGAGCAAGAGAC	303	
DB	181	AACCTGCACCTGTACCAATGCTCCCTGTGGCTGTCTCACCCTAGAGCAAGAGAC	240	
QY	304	AGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTGTCTG	363	
DB	241	AGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTGTCTG	300	
QY	364	GCCACAGCCGCTAGGAGTGGTGGCAGCCAGGTTGCTGTGTCAAGCTTGCACACTC	423	
DB	301	GCCACAGCCGCTAGGAGTGGTGGCAGCCAGGTTGCTGTGTCAAGCTTGCACACTC	360	
QY	424	GAAGACTTCTCTGACACCCGAGTCCAGATTGTGGAGGAGAACGGGACAGACGCTTT	483	
DB	361	GAAGACTTCTCTGACACCCGAGTCCAGATTGTGGAGGAGAACGGGACAGACGCTTT	420	
QY	484	GAATGCCACACCAAGGCTTCCAGCCGCCCATCATTTACTTTGGGAAAGGACAGGTTGACC	543	
DB	421	GAATGCCACACCAAGGCTTCCAGCCGCCCATCATTTACTTTGGGAAAGGACAGGTTGACC	480	
QY	544	GTGCTGAGAGGCCCCGGCTCATCATCTTTCCCAAGTGGCTCTCCAGATCTTAGATGTC	603	
DB	481	GTGCTGAGAGGCCCCGGCTCATCATCTTTCCCAAGTGGCTCTCCAGATCTTAGATGTC	540	

QY	604	CAGGACAGTGATGACAGGCTCTCTACCGCTGCGTGGCCACCAATTTCAGCCCGCCCAACGATTC	663	
DB	541	CAGGACAGTGATGACAGGCTCTCTACCGCTGCGTGGCCACCAATTTCAGCCCGCCCAACGATTC	600	
QY	664	AGCCAGGAGGCTCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACCAAGGGGGCAG	723	
DB	601	AGCCAGGAGGCTCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACCAAGGGGGCAG	660	
QY	724	GATGTGGTCAATTTGTGGCAGCCCCCAGAGAACCAACGGTAGTGTCTTGGACAGAACTTAGTG	783	
DB	661	GATGTGGTCAATTTGTGGCAGCCCCCAGAGAACCAACGGTAGTGTCTTGGACAGAACTTAGTG	720	
QY	784	ATGGAGTGGTGGCCCTCTGTGTGACCCACCCCTTTTGTGTCTCTGGGTCCGACAGGATGGA	843	
DB	721	ATGGAGTGGTGGCCCTCTGTGTGACCCACCCCTTTTGTGTCTCTGGGTCCGACAGGATGGA	780	
QY	844	AAGCCTATCTCAGCGATGTCAATCTTTGGGGCCGAGCAATCTACTATCTATGCCACAGCGG	903	
DB	781	AAGCCTATCTCAGCGATGTCAATCTTTGGGGCCGAGCAATCTACTATCTATGCCACAGCGG	840	
QY	904	CAGCCTCGGCACCTCTGGAGTCTATGTCTGCGGAGCAACAAAGCCCTCAAGCGTCACTTC	963	
DB	841	CAGCCTCGGCACCTCTGGAGTCTATGTCTGCGGAGCAACAAAGCCCTCAAGCGTCACTTC	900	
QY	964	GCACCTGGGGCTGTGAGCTCCGAGTGTCTGCTGCCCCAGCCCATCTCGCAGGACACCCGAG	1023	
DB	901	GCACCTGGGGCTGTGAGCTCCGAGTGTCTGCTGCCCCAGCCCATCTCGCAGGACACCCGAG	960	
QY	1024	GGCCTCTCGCGACCGCGGCGCAGACCGCGCGCTTTGTGTGCGGGGCTTCGGGGGAGCCA	1083	
DB	961	GGCCTCTCGCGACCGCGGCGCAGACCGCGCGCTTTGTGTGCGGGGCTTCGGGGGAGCCA	1020	
QY	1084	CGGCGCGGCTGCACCTGGCTGCACACCGGATCCCGTTGCGACCAATGGGGCGCTCAAG	1143	
DB	1021	CGGCGCGGCTGCACCTGGCTGCACACCGGATCCCGTTGCGACCAATGGGGCGCTCAAG	1080	
QY	1144	GTGCAAGGCGGTGGGGCGCAGCTTGTGTCTCATCTCAGATCGGCTCTGAGGACGCTTGGCTAC	1203	
DB	1081	GTGCAAGGCGGTGGGGCGCAGCTTGTGTCTCATCTCAGATCGGCTCTGAGGACGCTTGGCTAC	1140	
QY	1204	TACAGGTGCTAGCAGAAAAAAGACGCGGGGAACTGCTGTGCGGCTCGCGCCCTTGGCGGTA	1263	
DB	1141	TACAGGTGCTAGCAGAAAAAAGACGCGGGGAACTGCTGTGCGGCTCGCGCCCTTGGCGGTA	1200	
QY	1264	GTGGTGGCGAGGGGCTGCCAGCGCCCGCTGCGGCTCAGACGACGCGCTGAGCAGC	1323	
DB	1201	GTGGTGGCGAGGGGCTGCCAGCGCCCGCTGCGGCTCAGACGACGCGCTGAGCAGC	1260	
QY	1324	TCCTCTGTGCTGGTGGCGGCTGAGTTGTCACAGCGAGCAAAATCATTTGGCTTC	1383	
DB	1261	TCCTCTGTGCTGGTGGCGGCTGAGTTGTCACAGCGAGCAAAATCATTTGGCTTC	1320	
QY	1384	TCTCTTCACTACAAAAGGCAAGGGAGTGGACAATGTGGAGTACAGTTTGCAGTAAAC	1443	
DB	1321	TCTCTTCACTACAAAAGGCAAGGGAGTGGACAATGTGGAGTACAGTTTGCAGTAAAC	1380	
QY	1444	AAATGACACACAGAGCTGAGGTTGGGACCTGGACCCCAACAGGATTTAGTTCTAC	1503	
DB	1381	AAATGACACACAGAGCTGAGGTTGGGACCTGGACCCCAACAGGATTTAGTTCTAC	1440	
QY	1504	GTGGTGGCTTACTCCAGCTGGGGCGCAGCCGAACTCCAGCCAGCCCTGGTGCATACA	1563	
DB	1441	GTGGTGGCTTACTCCAGCTGGGGCGCAGCCGAACTCCAGCCAGCCCTGGTGCATACA	1500	
QY	1564	CTGGAAGATGTCCCGACGACGACCCAGCTTACCTGTTCAGGCCCAACCCCTCGGAC	1623	
DB	1501	CTGGAAGATGTCCCGACGACGACCCAGCTTACCTGTTCAGGCCCAACCCCTCGGAC	1560	
QY	1624	ATCAGGGTGGCATGGCTGCCCTTCCAGCTGAGCAATGGACAGGTGTGAGTAC	1683	
DB	1561	ATCAGGGTGGCATGGCTGCCCTTCCAGCTGAGCAATGGACAGGTGTGAGTAC	1620	
QY	1684	AAGATAGATGACGTTTGGGGAGGAAGATCAGGTTTTCCTCCACCGAGTGTCTGGAAT	1743	

1621 AAGATAGAGTACGGTTTGGGGAAGGAAGATCAGGTTTCTCCACCGAGGTGCTCGAAAT 1680
1744 GAGACAACTTACGTTAACTCCTTACAGCAAAACAAAGTGTACCGAGTCCGATTTCA 1803
1681 GAGACAACTTACGTTAACTCCTTACAGCAAAACAAAGTGTACCGAGTCCGATTTCA 1740
1804 GCTGACCTGCGCTGGCTTATGAGTCCCTTCTCAGTGGATGACGACACAGACACCTGGT 1863
1741 GCTGACCTGCGCTGGCTTATGAGTCCCTTCTCAGTGGATGACGACACAGACACCTGGT 1800
1864 GTGCAACACAGACGATGTTCCCTTGGCCCTCAGAAATGAAAGTGTAGGCGCAAGATG 1923
1801 GTGCAACACAGACGATGTTCCCTTGGCCCTCAGAAATGAAAGTGTAGGCGCAAGATG 1860
1924 GAGTCCCTGGTGTGTCATGGCAGCGCCCTCACCACCCACGATCTCTGGATACAA 1983
1861 GAGTCCCTGGTGTGTCATGGCAGCGCCCTCACCACCCACGATCTCTGGATACAA 1920
1984 CTCTACTGGGAGAGGTGGAAACAGAGGAGGAGGATGTTGACCGCCCCCAGGGGGT 2043
1921 CTCTACTGGGAGAGGTGGAAACAGAGGAGGAGGATGTTGACCGCCCCCAGGGGGT 1980
2044 COTGAGATCAAGCTTGGGACGTGGGCGCGCTGCGGCTGAAGAAAGTGAAGCAGTAT 2103
1981 COTGAGATCAAGCTTGGGACGTGGGCGCGCTGCGGCTGAAGAAAGTGAAGCAGTAT 2040
2104 GAACTGACCCAGTGTAGTCCCTGGCAGCGCTGACGAGTGAAGTGTGAGTCTTCAACAA 2163
2041 GAACTGACCCAGTGTAGTCCCTGGCAGCGCTGACGAGTGAAGTGTGAGTCTTCAACAA 2100
2164 CACGAGGAGCGTCTGCTGTGTGGAAGGCAAGACGAGAGCGCCACGCGACAC 2223
2101 CACGAGGAGCGTCTGCTGTGTGGAAGGCAAGACGAGAGCGCCACGCGACAC 2160
2224 CTGCTATCCAGAGGGGTCACCGTGTCTCTGCTCCATGTCCACGAGTCAACACG 2283
2161 CTGCTATCCAGAGGGGTCACCGTGTCTCTGCTCCATGTCCACGAGTCAACACG 2220
2284 TCACCTTCAATTTGGCTTGGTGGGAAGACGACAGCTTACACCTGTCAGATTGTCAAC 2343
2221 TCACCTTCAATTTGGCTTGGTGGGAAGACGACAGCTTACACCTGTCAGATTGTCAAC 2280
2344 TACACTGTACGCTTCGGCCCTCGGGGCTCAGAAATGCTTCCCTGCTACCTTACTATACC 2403
2281 TACACTGTACGCTTCGGCCCTCGGGGCTCAGAAATGCTTCCCTGCTACCTTACTATACC 2340
2404 AGCTCTGGAGAAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACAGTTTGG 2463
2341 AGCTCTGGAGAAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACAGTTTGG 2400
2464 GTACAGTCCACGAGTGTATGATGAGTGGGCCCTTTGGCTCGCTGTAGAACGCTCCACC 2523
2401 GTACAGTCCACGAGTGTATGATGAGTGGGCCCTTTGGCTCGCTGTAGAACGCTCCACC 2460
2524 CTGCTGACCGGCTTCAACCTCTCTGACCTGCGCTGAGCGCCCTGAGCCCTGACACCTCC 2583
2461 CTGCTGACCGGCTTCAACCTCTCTGACCTGCGCTGAGCGCCCTGAGCCCTGACACCTCC 2520
2584 ACCGTTGGGTTACACTGGTGTCCCTCCACGAGGCAATGTTGAGATTGTGAGTATCTA 2643
2521 ACCGTTGGGTTACACTGGTGTCCCTCCACGAGGCAATGTTGAGATTGTGAGTATCTA 2580
2644 ATTCTTACAGCAACAAACACACCGAGCCGCAACACAGTGTGACACTGTCTACACAGAG 2703
2581 ATTCTTACAGCAACAAACACACCGAGCCGCAACACAGTGTGACACTGTCTACACAGAG 2640
2704 GGAACATCTTCAGTGCAGAGTTCATGGCTTACAGAGTGAACACTCGTATTTCTCAAG 2763
2641 GGAACATCTTCAGTGCAGAGTTCATGGCTTACAGAGTGAACACTCGTATTTCTCAAG 2700
2764 ATGGAGGCGCCACAGAGTGGGCTGGGCGCTTTTCCCGCTTGCAGGATGTGATTACT 2823

Db 2701 ATGGAGCCCCCAGACAGAGGTGGGCGCTTTCCCGCTTGCAGGATGTGATTACT 2760
QY 2824 CTGCAAGAGACATTTCTCAGACTCTCTTGATGTGAC 2859
Db 2761 CTGCAAGAGACATTTCTCAGACTCTCTTGATGTGAC 2796
RESULT 9
AC112161/c
LOCUS Mus Musculus chromosome 9 BAC clone MGSI-117K9 ES cell line,
DEFINITION Complete Sequence, complete sequence.
AC112161
VERSION AC112161.9 GI:21426127
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 115135)
Montgomery, K.T., Grills, G., Li, L., Brown, W.A., Chiu, D., Decker, J.,
Fusina, M., Haider, A., Keller, A., Perera, A., Shim, C., Thomas, E.,
Zenneck, W., Xi, C., Juels, P. and Kucherlapati, R.
Zenbeck, W., Xi, C., Juels, P. and Kucherlapati, R.
High Throughput Mouse Sequencing
Unpublished
2 (bases 1 to 115135)
Montgomery, K.T., Grills, G., Li, L., Brown, W.A., Chiu, D., Decker, J.,
Fusina, M., Haider, A., Keller, A., Perera, A., Shim, C., Thomas, E.,
Zenneck, W., Xi, C., Juels, P. and Kucherlapati, R.
Direct Submission
Submitted (20-FEB-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
3 (bases 1 to 115135)
Montgomery, K.T., Grills, G., Li, L., Brown, W.A., Chiu, D., Decker, J.,
Fusina, M., Haider, A., Keller, A., Perera, A., Shim, C., Thomas, E.,
Zenneck, W., Xi, C., Juels, P. and Kucherlapati, R.
Direct Submission
Submitted (08-JUN-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
4 (bases 1 to 115135)
Montgomery, K.T., Grills, G., Li, L., Brown, W.A., Chiu, D., Decker, J.,
Fusina, M., Haider, A., Keller, A., Perera, A., Shim, C., Thomas, E.,
Zenneck, W., Xi, C., Juels, P. and Kucherlapati, R.
Direct Submission
Submitted (14-JUN-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
5 (bases 1 to 115135)
Montgomery, K.T., Grills, G., Li, L., Brown, W.A., Chiu, D., Decker, J.,
Fusina, M., Haider, A., Keller, A., Perera, A., Shim, C., Thomas, E.,
Zenneck, W., Xi, C., Juels, P. and Kucherlapati, R.
Direct Submission
Submitted (21-SEP-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
On Jun 14, 2002 this sequence version replaced gi:21358700.
-----Genome Center:
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site: <http://www.hpcgg.org/Sequence/mouse.html>
Contact: hpgc@mendel.mgh.harvard.edu

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550).
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double stranded sequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation as Low Coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as Low Quality.

-----Summary Statistics

Center project name: ALH
Sequencing vector: pSMART; AF399742
Chemistry: Dye-terminator Big Dye; 100%
Assembly program: Phrap version 0.990319
Contig length: 115035
Fraction of Phrap value < 40: 0
Error Rate in Consed: 0.00 per 10,000 bases
Number of N's in consensus: 0

----- Distribution of Quality < 40 Bases:

# bases	5	10	15	20	25	30	35	40
1000								
900								
800								
700								
600								
500								
400								
300								
200								
100								
0								

FEATURES

source	Location/Qualifiers
repeat_region	1. .115135
repeat_region	/organism="Mus musculus"
repeat_region	/mol_type="genomic DNA"
repeat_region	/strain="ES cell line"
repeat_region	/db_xref="taxon:10090"
repeat_region	/chromosome="9"
repeat_region	/clone="MGS1-117K9"
repeat_region	190..274
repeat_region	/rpt_family="ID_B1"
repeat_region	275..497
repeat_region	/rpt_family="B3"
repeat_region	498..652
repeat_region	/rpt_family="ID_B1"
repeat_region	760..829
repeat_region	/rpt_family="L2"
repeat_region	870..990
repeat_region	/rpt_family="ORR1D"
repeat_region	1173..1422
repeat_region	/standard_name="C79952"
repeat_region	2880..2914
repeat_region	/rpt_family="(TG)n"
repeat_region	complement(2918..3013)
repeat_region	/rpt_family="RSINE1"
repeat_region	3119..3173

Query Match 44.4%; Score 2743.2; DB 10; Length 115135;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2818; Conservative 0; Mismatches 13; Indels 7; Gaps 6;
QY 3341 AGAGCAACGGGAGAGAGCGCTCCAGCCTCGAGGAATCAGGTGGAGCTGAGGTCA 3400
Db 105112 AGAGCAACGGGAGAGAGAGCGCTCCAGCCTCGAGGAATCAGGTGGAGCTGAGGTCA 105053
QY 3401 TTGTCCACTCCGACTTCGGTGCATCCAAAGGATGCTCCTGACCTCCACCTCCAGACCTGG 3460

Db	105052	TTGTCACCTCCGACTTGGTGGCATCCAAAGGATGCTCTGACCTCCACCTCCAAAGACCTGG	104993
Qy	3461	AGCCAGAGAAACCACTGACTGACAGAGACTCTGGCTTCCACAGTCTGGAGCTGGAGTCTGT	3520
Db	104992	AGCCAGAGAAACCACTGACTGACAGAGACTCTGGCTTCCACAGTCTGGAGCTGGAGTCTGT	104933
Qy	3521	CTCAAGGAGCAGACTGGCTGGGACAGGAGCTGGGAGGTCGCAACCAACCAAGTGGGC	3580
Db	104932	CTCAAGGAGCAGACTGGCTGGGACAGGAGCTGGGAGGTCGCAACCAACCAAGTGGGC	104873
Qy	3581	CAGAGAGCTCACCCTGCTGGCAGAGCAGCAGTGCCTCTGCTCTGCTCAGACCTCC	3640
Db	104872	CAGAGAGCTCACCCTGCTGGCAGAGCAGCAGTGCCTCTGCTCTGCTCAGACCTCC	104813
Qy	3641	AGCCAGACACTGCTATAGAGGAGCCCTGGGAAAGCTGCAGGCCCAAGCCCTGTGTC	3700
Db	104812	AGCCAGACACTGCTATAGAGGAGCCCTGGGAAAGCTGCAGGCCCAAGCCCTGTGTC	104753
Qy	3701	CTCTAACAGTCAAGCCCAAGCCTTCCAGGGCCCTGTCTCTCTGCTCAGGTCCCTGAG	3760
Db	104752	CTCTAACAGTCAAGCCCAAGCCTTCCAGGGCCCTGTCTCTCTGCTCAGGTCCCTGAG	104693
Qy	3761	CAGAAAGCAGATATGGCTCAGGAAATGCGCATGCGATGGGTACATGTGTGTACAGAGA	3820
Db	104692	CAGAAAGCAGATATGGCTCAGGAAATGCGCATGCGGTACATGTGTGTGTACAGAGA	104633
Qy	3821	TATCCATAGTCTTGGAGCCTCTTAGGCTCTTGGCTGGGTTGGGGAGAACCTTTACT	3880
Db	104632	TATCCATAGTCTTGGAGCCTCTTAGGCTCTTGGCTGGGTTGGGGAGAACCTTTACT	104573
Qy	3881	CTCCCTCATATTTCTGCATCACAACAGGAGGACTTGGAGACACAGCTCTGTGTATGGAC	3940
Db	104572	CTCCCTCATATTTCTGCATCACAACAGGAGGACTTGGAGACACAGCTCTGTGTATGGAC	104513
Qy	3941	AGGTGTGAAGTCTGT	4000
Db	104512	AGGTGTGAAGTCTGT	104453
Qy	4001	CCTATGTAGCACTACTGTGGCTAGTTGACCTCCGCGGAGGATGGTGAACAGTGTAT	4060
Db	104452	CCTATGTAGCACTACTGTGGCTAGTTGACCTCCGCGGAGGATGGTGAACAGTGTAT	104393
Qy	4061	CAGTGCCAGCTCTTTAGCTTTTAGCTTTGTACCTAGCTTTTATACCTCTGAGAGT	4120
Db	104392	CAGTGCCAGCTCTTTAGCTTTTAGCTTTGTACCTAGCTTTTATACCTCTGAGAGT	104333
Qy	4121	GTCTCCAGTGTGTGTCTACAAAGACAGGCGCCAGCTCTTCTGTGAGCTGTGTGAGC	4180
Db	104332	GTCTCCAGTGTGTGTCTACAAAGACAGGCGCCAGCTCTTCTGTGAGCTGTGTGAGC	104273
Qy	4181	AGAGTGCCAGTCAACTCCAGGCGCTATGACACCGCAGCTACACAGCATGGCTGTCTAT	4240
Db	104272	AGAGTGCCAGTCAACTCCAGGCGCTATGACACCGCAGCTACACAGCATGGCTGTCTAT	104213
Qy	4241	CCCCCTGGCTCCTTAAGGTCAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTT	4300
Db	104212	CCCCCTGGCTCCTTAAGGTCAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTT	104153
Qy	4301	GAGCATCTGTGACCTAATTTTGTAACTTGGGAAGTGTGGTTTGGGAATCTTCTTTTC	4360
Db	104152	GAGCATCTGTGACCTAATTTTGTAACTTGGGAAGTGTGGTTTGGGAATCTTCTTTTC	104093
Qy	4361	GCACCTGTCCCTCTCTGCCCCCTCTGCTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTG	4420
Db	104092	GCACCTGTCCCTCTCTGCCCCCTCTGCTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTG	104033
Qy	4421	TGCTGATATCTGCGGCGCTCTCTCTTCCCATGATGCGCCCTGATTTCTCAGTCTGT	4480
Db	104032	TGCTGATATCTGCGGCGCTCTCTCTTCCCATGATGCGCCCTGATTTCTCAGTCTGT	103973
Qy	4481	TTTTCATTTCTGTCTGCGCATGCTGTCTTTATGTCTGTGTGTGTGTGTGTGTGTGTGTGT	4540
Db	103972	TTTTCATTTCTGTCTGCGCATGCTGTCTTTATGTCTGTGTGTGTGTGTGTGTGTGTGTGT	103913
Qy	4541	CTATGACACCTTTCTTAACCAACATGACTACCTCATGTCTGTCTCAGACCACTAGTGTGACC	4600
Db	103912	CTATGACACCTTTCTTAACCAACATGACTACCTCATGTCTGTCTCAGACCACTAGTGTGACC	103853
Qy	4601	CCTGGGTCCCAACAGCTCTCCCTGCAACCGCTTCTCTGGGAGATGAGCCCATCTCCAACT	4660
Db	103852	CCTGGGTCCCAACAGCTCTCCCTGCAACCGCTTCTCTGGGAGATGAGCCCATCTCCAACT	103794
Qy	4661	AGATCTGGAAGAACCTTGTGGCTTCTGTGGCTCTGCTCTCCCTCTGGTGTGAGATGAGA	4720
Db	103793	AGATCTGGAAGAACCTTGTGGCTTCTGTGGCTCTGCTCTCCCTCTGGTGTGAGATGAGA	103735
Qy	4721	AGGTTTTCTATGGAAGAGATGAGTCCAGGCTGCAACAGGGGAACCCCAAGAGGGGTAGG	4780
Db	103734	AGGTTTTCTATGGAAGAGATGAGTCCAGGCTGCAACAGGGGAACCCCAAGAGGGGTAGG	103675
Qy	4781	GAGTGAACCAAGAGGCTGAAAAAATGCTGCAACCGCTGCAACAGAGAGATGGGTG	4840
Db	103674	GAGTGAACCAAGAGGCTGAAAAAATGCTGCAACCGCTGCAACAGAGAGATGGGTG	103615
Qy	4841	TGTGCTTTTGACGCTGCAAGTCTGCTGCTGAACTGAAAGGGGTGAGAGAGGGGTACTG	4900
Db	103614	TGTGCTTTTGACGCTGCAAGTCTGCTGCTGAACTGAAAGGGGTGAGAGAGGGGTACTG	103555
Qy	4901	GGCTGCATGGCTCAGTCTCCCTGACCTCGAGCCCTGAACTCTGGCTTCTCAGAGTACCAAG	4960
Db	103554	GGCTGCATGGCTCAGTCTCCCTGACCTCGAGCCCTGAACTCTGGCTTCTCAGAGTACCAAG	103495
Qy	4961	AGTTTCTCCAAAGATGCTGTAAGGAAAGTCTTTGCAATAGGAAAGGGCGCTGCTCAT	5020
Db	103494	AGTTTCTCCAAAGATGCTGTAAGGAAAGTCTTTGCAATAGGAAAGGGCGCTGCTCAT	103435
Qy	5021	TTATTTTATCTTTCTTTACACTGCAATCCCAAAATCATCTTACCAAAAGGGCCCAAGCTG	5080
Db	103434	TTATTTTATCTTTCTTTACACTGCAATCCCAAAATCATCTTACCAAAAGGGCCCAAGCTG	103375
Qy	5081	ACTGGTATTTCTCAGTCAACAGAGCCATGCACTCTCTGGTCTTCTCAGCTCAGTCAAT	5140
Db	103374	ACTGGTATTTCTCAGTCAACAGAGCCATGCACTCTCTGGTCTTCTCAGCTCAGTCAAT	103315
Qy	5141	TCCAGAAATGTTCAGATCCAGTGGCATCTGTGCTCTTGTGCTGCAATCTTTCTATTTCAAC	5200
Db	103314	TCCAGAAATGTTCAGATCCAGTGGCATCTGTGCTCTTGTGCTGCAATCTTTCTATTTCAAC	103255
Qy	5201	TGGCTGGCAGATCAAGTGTAACTCTGGCTTCTGGGCGCAAGTGAATAACCAAGTCTAT	5260
Db	103254	TGGCTGGCAGATCAAGTGTAACTCTGGCTTCTGGGCGCAAGTGAATAACCAAGTCTAT	103195
Qy	5261	TTTCCCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	5320
Db	103194	TTTCCCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	103135
Qy	5321	TGAAGGCTGTGTGTATTTAGCTTGTGTGCTCATGTGCTGACCCCAAGTGTGCTGCTTC	5380
Db	103134	TGAAGGCTGTGTGTATTTAGCTTGTGTGCTCATGTGCTGACCCCAAGTGTGCTGCTTC	103075
Qy	5381	CTTCTCTCCCTCTTTCAGGCGCTATGATAACACTAAAGATTTATTAAGTGTGCTGCTTC	5440
Db	103074	CTTCTCTCCCTCTTTCAGGCGCTATGATAACACTAAAGATTTATTAAGTGTGCTGCTTC	103015
Qy	5441	TATCTCATTTAAAGACAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5500
Db	103014	TATCTCATTTAAAGACAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	102955
Qy	5501	GCCAAACCAACCGTATGTTCTTCAATGCTCTGAAAGGTCAAGAGCTCATTTTGTGTTTTC	5560
Db	102954	GCCAAACCAACCGTATGTTCTTCAATGCTCTGAAAGGTCAAGAGCTCATTTTGTGTTTTC	102895
Qy	5561	TGGTTAGATTTCTTTCTCTGCTTGAATGAATAAACCGTTTTTAAACAGTAGGCTCTTA	5620
Db	102894	TGGTTAGATTTCTTTCTCTGCTTGAATGAATAAACCGTTTTTAAACAGTAGGCTCTTA	102835

QY 5621 GCATCACACACATAGTCACTTCTCATGTTCTTGTTTAAACAAGCACTTGGAGTTCGGGT 5680
 |||||
 Db 102834 GCATCACACACATAGTCACTTCTCATGTTCTTGTTTAAACAAGCACTTGGAGTTCGGGT 102777
 |||||
 QY 5681 TTAATAATTAATAGTCTCAATAGTACAGCAATTTATTAACCCATTAGCTGGGTGGAAATTTG- 5739
 |||||
 Db 102776 TTAATAATTAATAGTCTCAATAGTACAGCAATTTATTAACCCATTAGCTGGGTGGAAATTTGT 102717
 |||||
 QY 5740 TTCTCAAAAGCAATTAAGTAAATCTGGTATCTGCCTATCACTACACAGTTGATAGAA 5799
 |||||
 Db 102716 TTCTCAAAAGCAATTAAGTAAATCTGGTATCTGCCTATCACTACACAGTTGATAGAA 102657
 |||||
 QY 5800 AGTAGCCAGCAACTCCTAGCATTTATATATGATGGGCTTCTGAGTAACCTGGGAGGTGTTA 5859
 |||||
 Db 102656 AGTAGCCAGCAACTCCTAGCATTTATATATGATGGGCTTCTGAGTAACCTGGGAGGTGTTA 102597
 |||||
 QY 5860 GCTTTGTGACTTTGTAGCACCAGGTCTTATTAGGAAAGTCTGTGGCCCTTTTACAGGGCA 5919
 |||||
 Db 102596 GCTTTGTGACTTTGTAGCACCAGGTCTTATTAGGAAAGTCTGTGGCCCTTTTACAGGGCA 102537
 |||||
 QY 5920 TTAGTCCCTTTGTGCTTTGCCATGGATGCCCTTAAGTCTTTGAGTCTCATTTAAGAAAT 5979
 |||||
 Db 102536 TTAGTCCCTTTGTGCTTTGCCATGGATGCCCTTAAGTCTTTGAGTCTCATTTAAGAAAT 102477
 |||||
 QY 5980 CTTTTCCTCAAGCATGACAAAGTGTATCGCAACTTACATGCTCACTGCTTTACCTG-G 6038
 |||||
 Db 102476 CTTTTCCTCAAGCATGACAAAGTGTAT-GCAATCTTACATGCTCACTGCTTTACCTG 102418
 |||||
 QY 6039 CTTAGTTTGTGCTGGGTATTATTAATGTCACCTTTCCAGCATCATGCTTCCCTCTTACAAAT 6098
 |||||
 Db 102417 CTTAGTTTGTGCTGGGTATTATTAATGTCACCTTTCCAGCATCATGCTTCCCTCTTACAAAT 102358
 |||||
 QY 6099 ATGATATCTTTATGTTACACTAAGTGTGTATCATGATCTGTCCTCTTAAGAAATTA 6158
 |||||
 Db 102357 ATGATATCTTTATGTTACACTAAGTGTGTATCATGATCTGTCCTCTTAAGAAATTA 102298
 |||||
 QY 6159 ATAACTATTATTCAGAC 6176
 |||||
 Db 102297 ATAACTATTATTCAGAC 102280
 |||||

RESULT 10

AC110235/c
 LOCUS AC110235 158055 bp DNA linear HTG 15-MAY-2002
 DEFINITION Mus musculus clone RP23-100M12, WORKING DRAFT SEQUENCE, 17 ordered pieces.
 ACCESSION AC110235
 VERSION AC110235.5 GI:20800205
 KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 158055)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Mus musculus, clone RP23-100M12
 Unpublished
 2 (bases 1 to 158055)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faroo, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Melrim, J., Meneus, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., and Lander, E.

Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 158055)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faroo, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Melrim, J., Meneus, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (15-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 15, 2002 this sequence version replaced gi:20531900.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L22383

Center clone name: 100_M_12

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 153807 bases at least Q40

Consensus quality: 155134 bases at least Q30

Consensus quality: 155784 bases at least Q20

Insert size: 218000; agarose-fp

Insert size: 156455; sum-of-contigs

Quality coverage: 7.1 in Q20 bases; agarose-fp

Quality coverage: 10.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 17 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 6459: contig of 6459 bp in length

* 6460 6559: gap of 100 bp

* 6560 6667: contig of 108 bp in length

TITLE
JOURNALREFERENCE
AUTHORSTITLE
JOURNAL

COMMENT

FEATURES	source	Location/Qualifiers
*	6668	6767: gap of 100 bp
*	6768	7863: contig of 1096 bp in length
*	7864	7963: gap of 100 bp
*	7964	9426: contig of 1463 bp in length
*	9427	9526: gap of 100 bp
*	9527	11011: contig of 1485 bp in length
*	11012	11111: gap of 100 bp
*	11112	12741: contig of 1630 bp in length
*	12742	12841: gap of 100 bp
*	12842	14789: contig of 1948 bp in length
*	14790	14889: gap of 100 bp
*	14890	17826: contig of 2937 bp in length
*	17827	17926: gap of 100 bp
*	17927	21244: contig of 3318 bp in length
*	21245	21344: gap of 100 bp
*	21345	27413: contig of 6059 bp in length
*	27414	27513: gap of 100 bp
*	27514	36381: contig of 8888 bp in length
*	36382	36481: gap of 100 bp
*	36482	45104: contig of 8623 bp in length
*	45105	45204: gap of 100 bp
*	45205	58369: contig of 13165 bp in length
*	58370	58469: gap of 100 bp
*	58470	85782: contig of 27313 bp in length
*	85783	85882: gap of 100 bp
*	85883	112706: contig of 26824 bp in length
*	112707	112806: gap of 100 bp
*	112807	153927: contig of 41121 bp in length
*	153928	154027: gap of 100 bp
*	154028	158055: contig of 4028 bp in length.
FEATURES		
source		
1. .158055		
/organism="Mus musculus"		
/mol_type="genomic DNA"		
/db_xref="taxon:10090"		
/clone="RP23-100M12"		
/clone_lib="RPC1-23 Female Mouse BAC"		
1. .6459		
/note="assembly_fragment"		
clone end:SP6		
vector_side:left"		
6560. .6667		
/note="assembly_fragment"		
6768. .7863		
/note="assembly_fragment"		
7964. .9426		
/note="assembly_fragment"		
9527. .11011		
/note="assembly_fragment"		
11112. .12741		
/note="assembly_fragment"		
12842. .14789		
/note="assembly_fragment"		
14890. .17826		
/note="assembly_fragment"		
17927. .21244		
/note="assembly_fragment"		
21345. .27413		
/note="assembly_fragment"		
27514. .36381		
/note="assembly_fragment"		
36482. .45104		
/note="assembly_fragment"		
45205. .58369		
/note="assembly_fragment"		
58470. .85782		
/note="assembly_fragment"		
85883. .112706		
/note="assembly_fragment"		
112807. .153927		
/note="assembly_fragment"		
154028. .158055		
/note="assembly_fragment"		

Db 3773 TTCTTCTCTGCTATCTCTCACACCTCTCCAGACTATGTCATCTGTTCTCTCTGCTG 3832
Qy 4533 AGTTCAACCTATGACACCTTCTTCTTAAACAATGACTACCTCATGTCGCTTCA----- 4585
Db 3833 GGTTCNAACCTCTGATCTCTTCTTAAACAATGACTACCTCATGTCGCTTCTTCAAGGCC 3892
Qy 4586 --GACCATAGTGTGACCCCTGGGTCCCAAGCTCCCTGCGCAACCGCTTCTCT----- 4637
Db 3893 CGTGCCCTTCTGATATCGCGGCTGCGCGCACTCGCTGCGCATCTCTCTGCTCTCTCT 3952
Qy 4638 -----GGCAGATGAGCCCACTCCAGTAGATCTGGA 4669
Db 3953 CACTCAGTGTCTGCTTGCCTGCGCCCGCCAGCGAGCCCAACCGCCAGTGGCGGTGG 4012
Qy 4670 AAAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4729
Db 4013 AGRAGATCTTCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4070
Qy 4730 ATGGAAGAGATGAGTCCAGGCTGCAAGGGGAAACCCCAAGAGGGGTAGGGAGTGAAC 4789
Db 4071 ATGGATG---GATCCAAGTCAATAGAGGGGAATGTTTGAGACAGGGAGGGGGCTGTGAT 4126
Qy 4790 CAAGGCTGAAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4846
Db 4127 CCAGAGGCTCAGAAATAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4186
Qy 4847 TTTGACGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4906
Db 4187 GGAGATGTTGGGCTCTCTTCAGAGAGAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAG 4246
Qy 4907 CATGGCTCAGTTCCTGACCTGAGCCCTGAGCCCTGAGCTGCTGCTGCTGCTGCTGCTGCT 4965
Db 4247 CAGGAGGCCCTTCCACCTCTGAGGCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4306
Qy 4966 -----CCTCAAGATGCTGTAAGGAGTCTTTGATAGAGAGAGAGAGAGAGAGAGAG 5020
Db 4307 CTCACCTCTGCGGTGCTGCGAGAGAGTCTTTCAGAGAGAGAGAGAGAGAGAGAGAG 4364
Qy 5021 TTATTTATCTTCTTACCTGATCCCAAGTCT-----CTTACCAAGAG 5070
Db 4365 CAGTTGTTTCTTAAATGAACTCTCAAGTCAATTTCTGTTCACTGCTGCTGCTGCTGCTGCT 4424
Qy 5071 GCCAAGCTGACTGCTATTTCTGAGTCAAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5130
Db 4425 GACAAGCTTGACTCTATTTCTGCTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4484
Qy 5131 CTCAGTCAATGCTCAGAGATGCTCAGATCCAGTGGCAATGCTGCTGCTGCTGCTGCTGCTGCT 5190
Db 4485 CTCAGCCCTCTCATAGAGAGCATAGAAATGTTAGGGTCTTTACTCCCTA----- 4531
Qy 5191 CTATTTCAACCTGCTGCAATCAAGTGTAACTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 5250
Db 4532 --ATGATGCTGATGAGCATCAAGAGTAACTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 4589
Qy 5251 ACCAGTCTATTTCTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 5304
Db 4590 ACCAGTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4649
Qy 5305 GTGAGTGTAGCTT-----CTGAAAGCGTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTT 5355
Db 4650 GCAGTGTGGGCTCTCTCAACCTGAAAGCATCTGTTTATTTATTTATTTATTTATTTATTTATTTATTT 4709
Qy 5356 TCATGTTGA---CCCAACCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5410
Db 4710 TCATGTTGATGCTGACCTTACGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4769
Qy 5411 TAACACTAAAGATTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5470
Db 4770 TAACACTAAAGATTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4829
Qy 5471 ACTACTTCTATAGCATTTCAAGTGGCCAGCGCCCAACCACT--CGTATGTTTCTTCTTCTTCTTCTTCTTCT 5528

Db 4830 AGTATTTCTGTAGCATTCAGCGCTGCTGTGGCTAAACACCACTGTGTATGTTTCATCATTTG 4889
Qy 5529 CTCTGAAGGTCAAGAGCTCATTTTGTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 5581
Db 4890 CTCTGAAGGTCAAGAGCTCATTTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 4949
Qy 5582 -----TGCTTGAATGAATAAACCGTTTAAACAGTAGGCTCTTAGCATCACACCAC 5632
Db 4950 AAAAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5009
Qy 5633 ATAGTCAATTTCTCATGTTCTTGTGTTTAAACAGCACTGAGGTTCTGGGTTTAAATTAATA 5692
Db 5010 ATAGTCAATTTCTCATGTTCTTGTGTTTAAACAGCACT--GAGGTTCTGCTGTTTAAATTAATA 5067
Qy 5693 GCTGCAAAATGAGCAAAATTAATAACCAATTAAGGCTGGGTGCAAAATTTG--TTCTCAAAAGCA 5751
Db 5068 GCTGCAAAATGAGCAAAATTAATAACCAATTAAGGCTGGGTGCAAAATTTGTTTCTCAAAAGCA 5127
Qy 5752 AATAAGTAATAATCTGCTATCTGCTATTAACCTACAGTTGATAAGAAAGTAGCCAGAAC 5811
Db 5128 AATAAGTAATAATCTGCTATCTGCTATTAACCTACAGTTGATAAGAAAGTAGCCATTTTC 5187
Qy 5812 TCAGTACATTAATATGATTTGGGTTCTGAGTAACTGGGAGTGTAGCTTTGTGACTT 5871
Db 5188 TCAGTACATTAATATGATTTGGGTTCTGAGTAACTGGGAGTGTAGCTTTGTGACTT 5247
Qy 5872 TGAGGACCAAGGTTCT--TATTAGGAAAGTCTGTTGGGCTTTTACAGGCACTTAGTCCCTTT 5930
Db 5248 TGAGGCAAGTATTTTATTAGAAAGAACTATTTAGGCTTTTACAGGCACTTAATCCCTTT 5307
Qy 5931 GTGCTTTGCGATGCTGCTTAAAGTTCTTTG--GAGTCTCATTTAAAGAAATTTCTTTCTCG 5989
Db 5308 GTCACTTACATTTGATGCTTAAAGTTTCTGAGTCTCAATTAAGAAATTTCTTTCTTTCTG 5367
Qy 5990 AAGCATGCAAGTGTATGCGCAATCTTACATGCTACTCTGTTTACCTG--GCTTAGTTTGT 6048
Db 5368 ATGATGACAAAGTGT-----AATCAGTACTTGTCTCATTTATTTGCTGTTATTTAGTTTAT 5422
Qy 6049 GCTGGGTATTTAATTTGCACTTTCCAGCATCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6108
Db 5423 GCTGACTATTTAATTTATCTTCCAGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5482
Qy 6109 TTATTTGTTACATAAGGTGTTGA--TCATGATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6167
Db 5483 TTAGTGTAGCTAAGGCAATGATTCATGATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5542
Qy 6168 TTTCCAG 6174
Db 5543 TTTCCAG 5549

RESULT 12
AX418406
LOCUS
DEFINITION
Sequence 1 from Patent WO0206329.
ACCESSION
AX418406
VERSION
AX418406.1 GI:21523336
KEYWORDS
unidentified
unidentified
unclassified.
ORGANISM
REFERENCE
1
AUTHORS
Rastelli, L., Shimkets, R.A., Zerhusen, B., Malyankar, U.M. and Padigaru, M.
TITLE
Human polynucleotides and polypeptides encoded thereby
JOURNAL
Patent: WO 0206329-A 1 24-JAN-2002;
Curagen Corporation (US)
FEATURES
source
1. .3741
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="NOVX nucleic acid with homology to NOPE-like"

BASE COUNT		708 a		1209 c		1150 g		673 t		1		others	
ORIGIN													
Query Match		43.4%;		Score 2681.8;		DB 6;		Length 3741;					
Best Local Similarity		83.7%;		Pred. No. 0;									
Matches 3141;		Conservative		0;		Mismatches 583;		Indels 30;		Gaps 8;			
QY	3	GGCGGGGGG	CAACGGGGG	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT
DB	9	GGGGGACG	CGGGGCG	GGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT
QY	63	CGGGGAGCT	GGCAATG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG
DB	69	CGGGGAGCT	GCTGCTT	GGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG	CGGGGAG
QY	123	GCAAGTAT	CTCGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT
DB	129	GCAAGTAT	CTCGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT	CGGGGCT
QY	183	TGCTGGGCT	CCGACGAG	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT
DB	189	CGTGGACCC	CCGACGAG	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT	GGTGCAT
QY	243	GAACCTGC	ACCTGCT	ACCTGCT	ACCTGCT	ACCTGCT	ACCTGCT	ACCTGCT	ACCTGCT	ACCTGCT	ACCTGCT	ACCTGCT	ACCTGCT
DB	249	CACTTTAC	CTGCTGCT	CTGCTGCT	CTGCTGCT	CTGCTGCT	CTGCTGCT	CTGCTGCT	CTGCTGCT	CTGCTGCT	CTGCTGCT	CTGCTGCT	CTGCTGCT
QY	303	CAGCGATG	ATGAGGAG	CTTATG	CTTATG	CTTATG	CTTATG	CTTATG	CTTATG	CTTATG	CTTATG	CTTATG	CTTATG
DB	309	CAG ---	TGAGAGT	CACTG	CACTG	CACTG	CACTG	CACTG	CACTG	CACTG	CACTG	CACTG	CACTG
QY	363	GGCCCA	CAGCCCT	AGGAGT	GGGAGT	GGGAGT	GGGAGT	GGGAGT	GGGAGT	GGGAGT	GGGAGT	GGGAGT	GGGAGT
DB	366	AGCCCA	CGGCGCT	CGGCT	CGGCT	CGGCT	CGGCT	CGGCT	CGGCT	CGGCT	CGGCT	CGGCT	CGGCT
QY	423	CGAAGACT	TTCTCT	CTGCA	CCGAGT	CCGAGT	CCGAGT	CCGAGT	CCGAGT	CCGAGT	CCGAGT	CCGAGT	CCGAGT
DB	426	CGCAGACT	TTCTCT	CTGCA	CCGAGT	CCGAGT	CCGAGT	CCGAGT	CCGAGT	CCGAGT	CCGAGT	CCGAGT	CCGAGT
QY	483	TGAATG	CCACCA	CAAGGG	CTTCC	AGCCCT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT
DB	486	TGAGTG	CCACCA	CAAGGG	CTTCC	AGCCCT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT
QY	543	CTGTCCT	GAGAGG	CCGCT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT
DB	546	ATTGCT	GAGAGG	CCGCT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT
QY	603	CCAGGAC	AGTGC	AGGCT	CTTAC	CGCT	CGCT	CGCT	CGCT	CGCT	CGCT	CGCT	CGCT
DB	606	TCAGGAG	AGTGC	AGGCT	CTTAC	CGCT	CGCT	CGCT	CGCT	CGCT	CGCT	CGCT	CGCT
QY	663	CAGCAGG	AGGCT	CGCT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT
DB	666	CAGCAGG	AGGCT	CGCT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT	CACTAT
QY	723	GGATGCT	CACTAT	GGTGC	AGCCCA	AGCCCA	AGCCCA	AGCCCA	AGCCCA	AGCCCA	AGCCCA	AGCCCA	AGCCCA
DB	726	GGATGCT	CACTAT	GGTGC	AGCCCA	AGCCCA	AGCCCA	AGCCCA	AGCCCA	AGCCCA	AGCCCA	AGCCCA	AGCCCA
QY	783	GATGGAG	TGCTG	GGGCT	CTG	GGGCT	CTG	GGGCT	CTG	GGGCT	CTG	GGGCT	CTG
DB	786	GATGGAG	TGCTG	GGGCT	CTG	GGGCT	CTG	GGGCT	CTG	GGGCT	CTG	GGGCT	CTG
QY	843	AAAGCT	ATCTCA	CGGAT	GTCTAT	CGGAT	GTCTAT	CGGAT	GTCTAT	CGGAT	GTCTAT	CGGAT	GTCTAT
DB	843	AAAGCT	ATCTCA	CGGAT	GTCTAT	CGGAT	GTCTAT	CGGAT	GTCTAT	CGGAT	GTCTAT	CGGAT	GTCTAT
QY	903	GCAGCT	CGGACT	CTCTAT	GTCTG	CGGACT	CTCTAT	GTCTG	CGGACT	CTCTAT	GTCTG	CGGACT	CTCTAT
DB	903	GCAGCT	CGGACT	CTCTAT	GTCTG	CGGACT	CTCTAT	GTCTG	CGGACT	CTCTAT	GTCTG	CGGACT	CTCTAT
QY	963	CGCCACT	GGGCT	GCTG	AGCT	CCGAG	TGCTT ---	GCTG	AGCT	CCGAG	TGCTT ---	GCTG	AGCT
DB	963	CGCCACT	GGGCT	GCTG	AGCT	CCGAG	TGCTT ---	GCTG	AGCT	CCGAG	TGCTT ---	GCTG	AGCT

Db	98090	TATCAACAAGCCCTTTGAGGCTCTAGGATCCTTTGGCTGGGAGGAGAACTTTATT	98031	97085	AGAAGTAGTTACAGAGCTGCAGTGAAGTCCCTGACCTTGGAGCCTTGAACTGA	97026
Qy	3879	CTCTCCCTCATATTCGATCACAC-----ATACAGGAGGAGCTTGAGACACAGCTCTGT	3931	4941	CTGGCTTACAGTAGCAAGAGAGTTCTCTCAAGAGTGTGTAAAGGAGAGTCTTTGCATAGG	5000
Db	98030	CTCTCCCTCATATTTTCATACATATGTGATACAGGAGGAGCTTGAGACACAGCTGCAT	97971	97025	CTGGATTCAGGCTGGCAAGAGTTCTCTCAAGAGTGTGTAAAGGAGAGTCTTCGCACAGG	96966
Qy	3932	GTAATGAGCAGTGTGAAGT	3991	5001	AAAAGGGCGCTCGCTCATTTTATTTTATCTTTTCTTACACTGAATCCCAAAATCATCTT	5060
Db	97970	GTGATGACA-----GGT	97937	96965	AAAAGGGCGCTCGCTCATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	96906
Qy	3992	AAACCTCTCCCTATGTAGCACTCACTGTGGCTAGTTGACCTCCGTGGGAGAGATGTGT	4051	5061	ACCACAAAGGGCGCAAGCTGACTGTATTTCTGTAGTGTGTGTGTGTGTGTGTGTGTGT	5120
Db	97936	AAACCTCTCCCTATGTAGCACTCACTGGGCTGTGATTTGACCTCCATGGCAGGAGCTGT	97877	96905	ACCACATAGGGCGCAAGCTGATTTGGTATTTCTGTAGTGTGTGTGTGTGTGTGTGTGT	96846
Qy	4052	AACAGTATCAGTGCAGCTCTTTGAGCTTTTACCTTTGTACCTAGCTTTTATTTACAC	4111	5121	GGTTTCTCACTCAGTGCATGTCCAGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5180
Db	97876	ACAGTATCAGTGCAGCTCTTTGAGC-TTTAGCCTTTGACCTAGCTTTTATTTACAC	97818	96845	GAATTTCTCACTCAGCGGTGTTCAGGATCATC-GAATGTGGGATCTGTGTCTCTTTCT	96787
Qy	4112	TCTGAGAGTGTCTCAGTGTGTGTCTACAAAGACAGCGCCAGCCCTCTTCTGTCACT	4171	5181	GCACATCTTTCTATTTCAACTGGCTGGCAATCAAGTGTGTAACTCTGGCTTCTGGGCAA	5240
Db	97817	TCTGAGAGTGTCTCAGTGTGTGTCTACAAAG-GGTGCAGGCTCTTCTGTGTCACT	97759	96786	GCACATCTTTCTATTT-TACTGGCTGGCAACCAATCATTAATCTGTGGCTTCTGGGCAA	96728
Qy	4172	GTGCTGAGAGTGTCCAGTCAACTCCAAGGCTGTATGACACCGAGCTTACACAGCAT	4231	5241	GTTAGAAATTAACAGTCTATTTTC-----	5265
Db	97758	GTGCTGAGAGTGTCCAGTCAAA--CGGACAGGCTACGACACGGCA-----ACAGC	97708	96727	ATTAGAAATTAACAGTCTATCTCTCTATATCTTCTTCTTCTTCTTCTTCTTCTTCT	96668
Qy	4232	GGCTGTATCCCTCGGCTCTTAAGTTCAGAGTCTGGGTGAACCCAGCTCAGCTCCC	4291	5266	CTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	5324
Db	97707	GGCGTATCCCTCGGCTCTTAAGTTCAGAGTCTGGGTGAACCCAG-----	97658	96667	TCCT	96608
Qy	4292	CTCTCTCTTTGAGCATCTCTGATCAATTTTGTATCTGGGAAGTGTCTGTGTGGAAA	4351	5325	AGCGTCTGTGTATTTAGCGCTGTGTGTCTCAATGTGTGACCCCAACCAATTTCTTCT	5384
Db	97657	CTCTCTCTTTGAGCATCTCTGATCAATTTTGTATCTGGGAAGTGTCTGTGTGGAAA	97598	96607	A-CATCTGTGTATTAAGCGCTCATGTGTCTCATGTATGATGCCACCTACATTTTC--C	96552
Qy	4352	CTCTCTCTTTGAGCATCTCTGATCAATTTTGTATCTGGGAAGTGTCTGTGTGGAT	4404	5385	TCCTCTCTCTCTCTGAGCGCTATGATTAACACTAAAGATTTATTAATGTCTGTGTATC	5444
Db	97597	CTCTCTCTTTGAGCATCTCTGATCAATTTTGTATCTGGGAAGTGTCTGTGTGGAT	97538	96551	TCCTCTCTCTCTCTGAGCGCTATGATTAACACTAAAGATTTATTAATGTCTGTGTATC	96492
Qy	4405	CTGTCTCATCTCTGT	4464	5445	TCATTAAGACAGAGTGTCTCACTTTGAATCTTCTATAGCATTTCAAAAGTGGCCACG	5504
Db	97537	CTGTCTCATCTCTGT	97478	96491	TCGTAAAGACAGAGTGTCTCACTTTGAATCTTCTATAGCATTTCAAAAGTGGCCACG	96432
Qy	4465	ATTTCCTCATCTGT	4524	5505	ACACCCGCTATGTCTTCTCATTTGCTGAGGCTCAAGAGCTCTCATTTTGTCTTCTGT	5564
Db	97477	ATTTCCTCATCTGT	97418	96431	ACACCCGCTATGTCTTCTCATTTGCTGAGGCTCAAAAGTCTCATTTTGTCTTCTGT	96372
Qy	4525	CGTCCCTGAGTCAACCTATGCACTCTTCTTAAACAATGATCTTCTTCTTCTTCTT	4584	5565	TAGATT-----CTTTCTCTCTCTGCTTGAATGAATAAACCGTTTTAACAGTAGGCT	5616
Db	97417	CGTCCCTGAGTCAACCTATGCACTCTTCTTAAACAATGATCTTCTTCTTCTTCTT	97358	96371	TAGATTCCCCCACCACCCCAACCCCTTGGCTTGAATGAATAAACCGTTTTAACAGTAGGCT	96312
Qy	4585	AGACCATAGTGTGACCTCTGGTCCCAAGCTCTTCTTCCCAAGCTCTTCTTCTTCTT	4644	5617	CTTAGCATCACACCATAGTCTTCTCATTTGCTGAGGCTCAAGAGCTCTCATTTTGTCTT	5676
Db	97357	AGACCATAGTGTGACCTCTGGTCCCAAGCTCTTCTTCCCAAGCTCTTCTTCTTCTT	97298	96311	CTTAGCATCACACCATAGTCTTCTCATTTGCTGAGGCTCAAGAGCTCTCATTTTGTCTT	96254
Qy	4645	TGAGCCCACTTCAAGTAGTCTGAAAGACCTTTGTGGCTTGTCTGCTGCTCCCTCCCT	4704	5677	GGGTTTAAATTAATAGCTGCAAAATGAGACAAATTTATAACCCATTAGGCTGGGTGAAA	5736
Db	97297	TGA-TCCATCTCAGGTAGTCTGGAAGA-CTTTGTAGTCTTCTGCTGCTGCTCTCT	97240	96253	TGGTTTAAATTAACAGCTGCAAAATGAGACAAATTTATAACCCATTAGGCTGGGTGAAA	96194
Qy	4705	TGGTGTGAGATGAGAAGTTTCTATGGAAGATGAGTCCAGGCTGCAAGGGAAC	4764	5737	TTG-TTCTCAAAAGCAAAATAGTAATAAATCTGGTATCTGCTATATACTCAAGTTGATA	5795
Db	97239	TGGTATT-----GAGATGAACCCAGGCTCCCAAGGCTCCCAAGGCTCCCAAGGCT	97204	96193	TTGTTTCTCAAAAGCAAAATAGTAATAAATCTGGTATCTGCTATATACTCAAGTTGATA	96134
Qy	4765	CCCAAGAGGGGTAGGAGTGAACCAAGAGGCTGAAATAAATAATGGCTGCCACCCATCTG	4824	5796	AGAAAGTGGCGAAGTCTCACTAGCATTTATATATGATTTGGGTTCTGAGTAACTGGGAGT	5855
Db	97203	CTGAGAGGGGTAGGAGTGAACCAAGAGGCTGAAATAAATAATGGCTGCCACCCATCTG	97145	96133	AGAAAGTGGCGAAGTCTCACTAGCATTTATATATGATTTGGGTTCTGAGTAACTGGGAGT	96074
Qy	4825	CACAGAGAGTGGGTGTCTTTTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4884	5856	GTTAGCTTTTGTAGCTTTGTAGCACCGAGTCTTATTTAGGAA--AGTCTGTGGCTTTTAC	5913
Db	97144	CACAGAGAGTGGGTGTCTTTTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	97086	96073	GTTAGATTTGTGATTTTGTAGCAGCAGGTTTATTTAGGAAAGAGTCTATTTGGCTTTTAC	96014
Qy	4885	AGGGGAGCTACTGGGGCTGCCATGGCTCAGTTTCCCTGACCTTGGAGCTCTGAAC----	4940	5914	AGGGCATTTAGTCCCTTTGTCTTTGGCATGAGTCCCTTAAAGTTCTTTTGGAGTCTCATTTA	5973
				96013	AGGGTATTATCCCTTTGTCTTACCATGGATCCCTTAAAGTTCTTTTGGAGTCTCATTTA	95954


```
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11:349H15"
/clone_lib="RPC1-11 Human Male BAC"
1..1477
/note="assembly_fragment"
1578..2785
/note="assembly_fragment"
2886..4859
/note="assembly_fragment"
4960..6416
/note="assembly_fragment"
6517..8739
/note="assembly_fragment"
8840..10868
/note="assembly_fragment"
10969..13737
/note="assembly_fragment"
13838..15329
/note="assembly_fragment"
15430..17938
/note="assembly_fragment"
18039..19981
/note="assembly_fragment"
20082..23587
/note="assembly_fragment"
23688..27447
/note="assembly_fragment"
27548..30181
/note="assembly_fragment
clone_end:77
vector_side:left"
30262..34717
/note="assembly_fragment"
34838..40101
/note="assembly_fragment"
40202..44039
/note="assembly_fragment"
44140..49950
/note="assembly_fragment"
50051..55555
/note="assembly_fragment"
55656..60496
/note="assembly_fragment
clone_end:SP6
vector_side:right"
60597..68544
/note="assembly_fragment"
68645..76940
/note="assembly_fragment"
77041..85480
/note="assembly_fragment"
85581..93905
/note="assembly_fragment"
94006..108063
/note="assembly_fragment"
108164..123496
/note="assembly_fragment"
123597..139965
/note="assembly_fragment"
140066..176282
/note="assembly_fragment"
44272 a 43276 c 42765 g 43361 t 2608 others
ORIGIN
Query Match 13.8%; Score 854.8; DB 2; Length 176282;
Best Local Similarity 64.5%; Pred. No. 1.9e-193;
Matches 1930; Conservative 0; Mismatches 847; Indels 217; Gaps 35;
QY 3341 AGGCAACGGGAGAAAGACCGTCCCGAGCTGCAGGAATCAGGTGGAGCTGAGGTCA 3400
Db 154550 AGGGCAATGGGAGGAGAGTACACCCCGAGCTGCAGGAACAGGTGGAGGTGAGTCA 154491
```


Perreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kellis, C., LaRocque, K., Lamazares, R., Landers, J., Lechoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McSwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodores, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (23-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
3 (bases 1 to 198295)

AUTHORS
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., DeKrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kellis, C., LaRocque, K., Lamazares, R., Landers, J., Lechoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McSwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodores, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
4 (bases 1 to 198295)

AUTHORS
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., DeKrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kellis, C., LaRocque, K., Lamazares, R., Landers, J., Lechoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McSwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodores, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (26-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT
On Apr 26, 2002 this sequence version replaced gi:20148056.

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22971
Center clone name: 349_G_13
----- Location/Qualifiers
1. 198295
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-349G13"
/clone_lib="RPC1-11 Human Male BAC"
complement(1..79)
/rpt_family="AluS"
complement(183..494)
/rpt_family="AluSg"
complement(1216..1416)
/rpt_family="L1MB8"
complement(1417..1709)
/rpt_family="AluSg"
complement(1710..2115)
/rpt_family="L1MB8"
2116..2169
/rpt_family="(TTTAA)n"
2186..2234
/rpt_family="AluSg/x"
complement(2236..2512)
/rpt_family="AluJb"
complement(2514..2631)
/rpt_family="L1MB8"
complement(2632..2941)
/rpt_family="AluY"
complement(2942..3104)
/rpt_family="L1MB8"
3181..3464
/rpt_family="AluSx"
3465..3711
/rpt_family="AluJb"
3712..3788
/rpt_family="(TA)n"
complement(4661..4940)
/rpt_family="AluJb"
complement(4946..5106)
/rpt_family="FRAM"
complement(5878..6193)
/rpt_family="AluYb8"
7197..7490
/rpt_family="AluSx"
7925..8226
/rpt_family="AluY"
complement(8227..8312)
/rpt_family="L1MB3"
8313..8347
/rpt_family="AT_rich"
complement(8350..8617)
/rpt_family="AluSx"
complement(8621..9274)
/rpt_family="L1MB3"
9403..9445
/rpt_family="AT_rich"
9706..10007
/rpt_family="AluJb"
complement(10940..11000)
/rpt_family="L1MB5"

All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22971
Center clone name: 349_G_13
----- Location/Qualifiers
1. 198295
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-349G13"
/clone_lib="RPC1-11 Human Male BAC"
complement(1..79)
/rpt_family="AluS"
complement(183..494)
/rpt_family="AluSg"
complement(1216..1416)
/rpt_family="L1MB8"
complement(1417..1709)
/rpt_family="AluSg"
complement(1710..2115)
/rpt_family="L1MB8"
2116..2169
/rpt_family="(TTTAA)n"
2186..2234
/rpt_family="AluSg/x"
complement(2236..2512)
/rpt_family="AluJb"
complement(2514..2631)
/rpt_family="L1MB8"
complement(2632..2941)
/rpt_family="AluY"
complement(2942..3104)
/rpt_family="L1MB8"
3181..3464
/rpt_family="AluSx"
3465..3711
/rpt_family="AluJb"
3712..3788
/rpt_family="(TA)n"
complement(4661..4940)
/rpt_family="AluJb"
complement(4946..5106)
/rpt_family="FRAM"
complement(5878..6193)
/rpt_family="AluYb8"
7197..7490
/rpt_family="AluSx"
7925..8226
/rpt_family="AluY"
complement(8227..8312)
/rpt_family="L1MB3"
8313..8347
/rpt_family="AT_rich"
complement(8350..8617)
/rpt_family="AluSx"
complement(8621..9274)
/rpt_family="L1MB3"
9403..9445
/rpt_family="AT_rich"
9706..10007
/rpt_family="AluJb"
complement(10940..11000)
/rpt_family="L1MB5"

repeat_region	11001..11284	/rpt_family="AluJo"	
repeat_region	complement(11285..11366)	/rpt_family="L1MB5"	
repeat_region	complement(11641..11934)	/rpt_family="AluSg"	
repeat_region	complement(12071..12294)	/rpt_family="AluJb"	
repeat_region	complement(12410..12697)	/rpt_family="AluSx"	
repeat_region	complement(12708..12998)	/rpt_family="AluJb"	
repeat_region	13158..13478	/rpt_family="AluJo"	
repeat_region	13970..14093	/rpt_family="FLAM_C"	
repeat_region	14108..14412	/rpt_family="AluJo"	
repeat_region	complement(14813..15119)	/rpt_family="AluSx"	
repeat_region	complement(15130..15432)	/rpt_family="AluSg"	
repeat_region	complement(15532..15841)	/rpt_family="AluSx"	
repeat_region	complement(15906..16077)	/rpt_family="AluSg/x"	
repeat_region	complement(16079..16383)	/rpt_family="AluSg"	
repeat_region	16409..16481	/rpt_family="MERSA"	
repeat_region	16482..16779	/rpt_family="AluSg"	
repeat_region	16780..16874	/rpt_family="MERSA"	
repeat_region	complement(17247..17543)		
Query Match 13.8%; Score 854.8; DB 9; Length 198295;			
Best Local Similarity 64.5%; Pred. No. 1.9e-193;			
Matches 1930; Conservative 0; Mismatches 847; Indels 217; Gaps 35;			
QY	3341	AGAGCAACGGGAGAAAGACCGTCCCGAGCTCGAGGAATCAGGTGGAGCTGAGGTCA	3400
DB	101335	AGGGCAATGGGAGGAAAGTCAACCCAGCTCGAGGAACCGAGGTGGAGGTCA	101394
QY	3401	TTGTCCATCCGACTTCGTGATCCAGAGGATGCTCGACCTCCACCTCCAGACCTGG	3460
DB	101395	TTGTCCATCCGACTTCGTGATCCAGAGGATGCTCGACCTCCACCTCCAGACCTGG	101454
QY	3461	AGCCAGAGAACCACTGATCGCAGAGACTCTGCCTTCCACGCTCGAGCTGGGATCTGT	3520
DB	101455	AGCCTGAGACCCCTGCTCCAGAGCTCTCGATCTCATCTCGGCTGGGGATCCAG	101514
QY	3521	CTCAAGGACGACCTGGCTGGGAGGAGGAGCTGGGAGGTCGCAACCAACAGTGGGC	3580
DB	101515	GGCAGGGGGCAGCTGGCTGGGAGGAGGAGTGGGAGGCTGAGCTGGCAGCCCGGGC	101574
QY	3581	CAGAGAGGCTCACTCTGTCAGAGAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT	3640
DB	101575	CAGACAGACTTACCTCTGTCAGAGAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT	101634
QY	3641	AGCCAGCAGCTGTATAGAGGAGGCCCCCTGGGAAAAGCTGCCAGGCCAAAGCCCTGTGTC	3700
DB	101635	AGCCAGCAGGCTGTATAGAGGAGGCCCCCTGGGAAAAGCTGCCAGGCCAAAGCCCTGTGTC	101694
QY	3701	CTCTAAGCTAGCCCAAGCTTCCAGGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3760
DB	101695	CTCTAAGCTAGCCCAAGCTTCCAGGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	101754
QY	3761	CAGA-----AGCAGATGGCTCAGGA-----CATGCCA	3791
DB	101735	CAGAGGATGTGTTGGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	101814
QY	3792	TGCAATGGCTACACATGTGTACTAGAGATATCCATAAGTCTTGGAGCTCTTAGGGTC	3851

DB	101815	TGATGTCCACATGTGTGCTGTAGATACATCATCAAGCCCTTGGAGCTTCTTAAGTTG	101874
QY	3852	CTTTGGCTGGGTTGGGAGAA-----CTTTACTCTCCCTCATATTTCTGCATCACATAC	3905
DB	101875	CTTTGGCTGGGAGGAGGAGAAACATGGAATATTTCACTCCCCCCATATCTTTGTGATAC	101934
QY	3906	AGGAGGACCTTGAGACACAGCTCTGTGTATATGAGACAGTGTGGAAGTGTGTGTGTGTG	3965
DB	101935	ACATGTGACATGTGAAGACATACGACATAGCTACATGTGTGACATGTGTGTGAAG	101994
QY	3966	TGTGTGTGTG-----TG	4020
DB	101995	TGCATGTATGCTACTG	102054
QY	4021	GCCTAGTTGACCTCCCTGCGGAGGATGTG--TAACAGTGTATCAGT--GCCAGCTCTTTGA	4077
DB	102055	GCCTAGTTGCTCTCAGGTGAGAACGCTGCGGAGGATGTGTGTGTGTGTGTGTGTGTGTGTG	102114
QY	4078	GCCTTTAGCTTGTACCTAGCTTTTATTT--ACACTCTGAGAGTGTCTCCAGTGTGTGTG	4135
DB	102115	GCCTTTAGCTTGTACCTAGCTTTTATTTATACACTCTGAGAGTGTCTCCAAATGCTGTG	102174
QY	4136	TCTAAGAGACGCGCCAGCCCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	4195
DB	102175	CTGACAAAGACGCTCCAGCCCATTTCTC--CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	102225
QY	4196	TCCAGGGCTTATGACACCGGAGCTTACCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	4255
DB	102226	-----GAGGCTCTGAATGCTGGAATTTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTG	102266
QY	4256	AGGTCCAGATG	4315
DB	102267	AGTCTCTTATTTGCCCAATAGAGAGGTGGCCCTGGCTCCCTCCGAGCAACTCTGCAATT	102326
QY	4316	TAATTTGTAACTCTGGAGGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	4375
DB	102327	TAATTTGTAACTCTGGAGGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	102386
QY	4376	CTGCGCTTCTCTCATTTG	4433
DB	102387	CTTCTCTTGGCTTG	102446
QY	4434	TGGGCTTCTCTCTT--TCCCATGATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	4487
DB	102447	CAGTGGCTTCTCTGTCTCCAGATCTCTTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTC	102506
QY	4488	-----TCTG	4542
DB	102507	TGCTATCTCTACACTCTCCAGACTATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	102566
QY	4543	ATGACCTTCTTCTTAAACATGACTACTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	4593
DB	102567	CTGCATCTTCTTAAACATGACTACTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	102626
QY	4594	TGTGACCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG	4638
DB	102627	CTGTATCCGGCTGCTGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	102686
QY	4639	-----GGCAGATGAGCCCACTTCAAGTAGATCTCAGAAAGACCTTT	4679
DB	102687	TTTG	102746
QY	4680	GTGGCTTG	4739
DB	102747	CTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG	102801
QY	4740	TGAGTCCAGGCTGACAGGGGAAACCCCAAGAGGGGTAGGAGTGTGTGTGTGTGTGTGTGTG	4799
DB	102802	-GATCAAGTATAGAGGGGATG	102860
QY	4800	AAAAAAATGGCTGCCACCC--ATCTGCACAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTG	4856

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 17:04:36 ; Search time 14061.8 Seconds
(without alignments)
10674.595 Million cell updates/sec

Title: US-09-754-997A-1
Perfect score: 6176
Sequence: 1 atggcgggcgacacggg.....taataaattttccagac 6176

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_hcc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_mam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rod:*
 - 26: em_gss_phg:*
 - 27: em_gss_vrl:*
 - 28: gb_gss1:*
 - 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1479	23.9	2626	11	AK079073 Mus muscu
2	883	14.3	947	11	AK020142 Mus muscu
3	805.8	13.0	849	14	CD352600 UI-M-GIO-
4	792.2	12.8	817	14	CA317106 UI-M-FW0-

5	737.6	11.9	748	14	CD348071
6	726.6	11.8	741	14	CD352400
7	720.6	11.7	944	13	BU840589
8	714.8	11.6	808	13	BU705518
9	706.4	11.4	719	14	CA751384
10	703.8	11.4	740	13	BU701299
11	690	11.2	713	14	CB524810
12	685.8	11.1	699	14	CD351162
13	639.6	10.4	900	12	B1185886
14	629.4	10.2	631	14	CB518504
15	618.2	10.0	671	14	BY719391
16	617.2	10.0	624	13	BU056331
17	599.8	9.7	651	10	BB624536
18	582	9.4	591	13	BU057120
19	581.6	9.4	637	10	BB665894
20	580	9.4	630	10	BB626002
21	574	9.3	621	9	AW049847
22	573	9.3	660	10	BB071741
23	561	9.1	583	14	CA895918
24	560.4	9.1	632	10	BB649997
25	556.4	9.0	759	14	CB988160
26	555.4	9.0	579	28	AZ746941
27	548.8	8.9	553	14	CA316452
28	546.2	8.8	703	14	BY712424
29	541.6	8.8	584	28	AZ745366
30	538.4	8.7	1137	11	AK082162
31	534.4	8.7	544	14	CA892088
32	529	8.6	551	13	BU705885
33	513.2	8.3	578	14	CB607425
34	509	8.2	509	14	CA873325
35	508.8	8.2	723	14	CA510263
36	505	8.2	528	4	BX519858
37	503.6	8.2	717	13	BX100607
38	502	8.1	565	14	CA883712
39	501.6	8.1	542	4	BX523249
40	500.8	8.1	504	14	CA876082
41	491.6	8.0	553	14	CA871246
42	485.6	7.9	528	13	BU580111
43	462.4	7.5	809	29	BZ222348
44	453.4	7.3	455	14	CA883713
45	452	7.3	533	10	BB681084

ALIGNMENTS

RESULT 1
AK079073
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

AK079073 2626 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:9330155G14 product:hypothetical protein,
full insert sequence.
AK079073.1 GI:26098261
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

Db 795 AGTGATGAGTGGGCTCTGCTGACCCACCCCTTTTGTGCTCGGTCCGACAGGA 854
 Qy 840 TGGAAAGCCTATCTCCAGGATGTCTATCGTTCTGGGCGGACCAATCTACTCATCGCCAG 899
 Db 855 TGGAAAGCCTATCTCCAGGATGTCTATCGTTCTGGGCGGACCAATCTACTCATCGCCAG 914
 Qy 900 GCGGAGCTCGGCACTCTGGAGTCT-ATGTCGCGGAGCCAAAGCCCTCTCAGCGGTG 958
 Db 915 GCGCCAGGCTCGGCACTTTTGAATTTAAATTTTCCCAACCCACCAACCCCGCCAGCGCTT 974
 Qy 959 ACTTCGCCACCTCGGCGCTGCTGAGCTCGAGTGTCTGCTGCCCCCAGCCATCTCGAGG--C 1016
 Db 975 AATTTTCCCGCTTCCGCTTCTTAAGTTCCAATTTCTTTTCCCGCCACCTTTTCCAGGCC 1034
 Qy 1017 ACCGAGGCGCTCTCGCGGACCGGGCCAGCACCGCGCTTCTGTCGCGGCGCTCCGG 1076
 Db 1035 GCCCAAGCCCTTTTCCGNAACCGGCGCCACCCCGCTTCTTTTCCCGGCTTCGG 1094
 Qy 1077 GGAGCCAGCGCGCGCTGCACTGCTGTCAGCGGATCCGTTTGCAGCCCAATGGGCG 1136
 Db 1095 GGACCCCGCGCGCGCTGCACTGCTGTCAGCGGATCCGTTTGCAGCCCAATGGGCG 1154
 Qy 1137 CGTCAAGTGCAGGCGCGCTGCGGCGAGCTTGTCTCATCTCATGATCGGCTGCAGGACGC 1196
 Db 1155 CGTCAAGTGCAGGCGCGCTGCGGCGAGCTTGTCTCATCTCATGATCGGCTTTCAGGACGC 1214
 Qy 1197 TGGCTACTACGAGTGCCTAGCAGAAACAGCGCGGAACTGCTGTGCGCTGCGCCCT 1256
 Db 1215 TGGCTACTACGAGTGCCTAGCAGAAACAGCGCGGAACTGCTGTGCGCTGCGCCCT 1274
 Qy 1257 GCGGCTAGTGTGCGGAGGCGCTGCCAGCGCCCGACTCGGGTCACAGCCACCGCCCT 1316
 Db 1275 GCGGCTAGTGTGCGGAGGCGCTGCCAGCGCCCGACTCGGGTCACAGCCACCGCCCT 1334
 Qy 1317 GAGCAGCTCTCTGCTGCTGCTGCTGAGGAGCGGCTGAG-TTGCAGACGCGGAAATCA 1375
 Db 1335 GAGCAGCTCTCTGCTGCTGCTGCTGAGGAGCGGCTGAG-TTGCAGACGCGGAAATCA 1394
 Qy 1376 TTGGCTT-CTCTCTTCACTACCAAGGAGGAGGAGTGGACATG--TGAGTACACG 1431
 Db 1395 TTGGCTTCTCTCTTCACTACCAAGGAGGAGGAGTGGACATG--TGAGTACACG 1454
 Qy 1432 TTTCAGTAAACCAATGACACACAG------GCTGCAGGTTCG 1469
 Db 1455 TTTCAGTAAACCAATGAGGAGGAGGAGTGGAGGCTTTCGGGACCTCG 1514
 Qy 1470 GGACTGGAAACCAACAGGATATGATGTTCTACGTGTGCTGCTACTCCAGCTGGGCGC 1529
 Db 1515 GAAACCCCAACCCCGATTAATGAAGTTTAAAGTGTGGCTTAACCTTCCAAAGCTGGCG 1574
 Qy 1530 CAGCC-----GAACCTCCAGCCAG--CCCTGGTGCATACACTGGACGATGTCGCCAG 1580
 Db 1575 GGGCCAGGCGGAAACCTTCAAGCCCGGCGCTGCTGTCATACACTGGACGATGTCGCCAG 1634
 Qy 1581 CGCAGCACCCAGCTTACTCTGTCAGCCCAACCCCTCGGACATC-AGGGTGGCATGCG 1639
 Db 1635 CGCACACCCAGCTTACTCTGTCAGCCCAACCCCTCGGACATC-AGGGTGGCATGCG 1694
 Qy 1640 TGCCCTCGCTCCAGCTGAGCAATGGAAGTGTCTGAAGTACAGATAGAGTACGGTT 1699
 Db 1695 TGCCCTCGCTCCAGCTGAGCAATGGAAGTGTCTGAAGTACAGATAGAGTACGGTT 1754
 Qy 1700 TGGGAAG--GAAATGAGTTTCTCAGCGAGTGTCTGGAATGAGACACAACTTA 1756
 Db 1755 TGGGAAGGAAAGATGAGTTTCTCAGCGAGTGTCTGGAATGAGACACAACTTA 1814
 Qy 1757 CGTTAACTCACTCAGCAACAAAGTGTACCGAGTCCGATTTTCAGTGCCTGCGG 1816
 Db 1815 CGTTAACTCACTCAGCAACAAAGTGTACCGAGTCCGATTTTCAGTGCCTGCGG 1874
 Qy 1817 CT-GGCTATGGAGTCCCTTCTCAGTGGATGACAGCAGCAGCTGCTGTCACAAACAG 1875

Db 1875 CTGGGCTATGGAGTCCCTTCTCAGTGGATGCAGCAGCAGCAGCTGCTGGTGGACACCG 1934
 Qy 1876 AGCCATGTT 1884
 Db 1935 AGCCATGTT 1943
 RESULT 2
 AK020142 947 bp mRNA linear HTC 05-DEC-2002
 LOCUS Mus musculus 12 days embryo male wolffian duct includes surrounding
 DEFINITION region cDNA, RIKEN full-length enriched library, clone:6720465A10
 product:hypothetical protein, full insert sequence.
 AK020142
 VERSION AK020142.1 GI:12860634
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 CARNINCI, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Mech. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 2
 CARNINCI, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
 Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
 Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
 Wyshaw-Borja, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
 and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085860
 PUBMED 11217851
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research

Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouseefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers

1. 849

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6856471"

/tissue_type="whole brain"

/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP G10"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;

Site 2: Not I; The library was constructed according

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with Ecor I adaptor, digested with Not I and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is ACGGAGACAG. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator." 237 t 7 others

BASE COUNT 176 a 225 c 204 g 237 t 7 others

Query Match 13.0%; Score 805.8; DB 14; Length 849;

Best Local Similarity 98.6%; Pred. No. 3.5e-125;

Matches 840; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

QY 4390 ATTGTTCTGGTGATCTGTCTCATGTCATCTTGTGCGATTATCTCGGGGCCCTTCTCTTT 4449

DB 1 ATTTGTTCTGGTGATCTGTCTCATGTCATCTTGTGCGATTATCTCGGGGCCCTTCTCTTT 60

QY 4450 CCATGATGCCCTGATTCCTCACTGCTGTTTTTCAATTTCTGTGCGCATGTTGCTCTTT 4509

DB 61 CCCATGATGCCCTGATTCCTCACTGCTGTTTTTCAATTTCTGTGCGCATGTTGCTCTTT 120

QY 4510 ATGTCGTGTTTCTCGTCCCTGAGTTCACCTATGCAACCTTCCCTTAACACATGACTA 4569

DB 121 ATGTCGTGTTTCTCGTCCCTGAGTTCACCTATGCAACCTTCCCTTAACACATGACTA 180

QY 4570 CCTCATGTCGTCTCAGACATAGTGTGACCCCTCGGTCGCCACAGCTCCCTCGCAACC 4629

DB 181 CTTGATGTCGTCTCAGACATAGTGTGACCCCTCGGTCGCCACAGCTCCCTCGCAACC 240

QY 4630 GCTTTCTGGGCGATGAGCCCACTCCAAGTATGATCTGGAAGAAGCCCTTGTGGCTTGT 4689

DB 241 GCTTTCTGGGCGATGAGCCCACTCCAAGTATGATCTGGAAGAAGCCCTTGTGGCTTGT 298

QY 4690 TGGCTGCCCTCCCTGTTGATGATGAGAGTTTCTATGGAAGATGATGATGATGATGAT 4749

DB 299 TGGCTGCCCTCCCTGTTGATGATGAGAGTTTCTATGGAAGATGATGATGATGATGAT 358

QY 4750 CTGCACAGGGGAACCCCAAGAGGGGTAGGAGTGAACCAAGAGGCTGAAAAAATG 4809

DB 359 CTGCACAGGGGAACCCCAAGAGGGGTAGGAGTGAACCAAGAGGCTGAAAAAATG 418

QY 4810 GCTGCCACCCATCTGCACAGAGATGGGTGTGTCTTTTGTACGTGCGATCTCGGCTGAA 4869

DB 419 GCTGCCACCCATCTGCACAGAGATGGGTGTGTCTTTTGTACGTGCGATCTCGGCTGAA 478

QY 4870 ACTGAAGGGGTGAGGAGAGGGGAGCTACTTGGGGCTGCCATGGCTCAGTTCCTGACCCCTG 4929

DB 479 ACTGAAGGGGTGAGGAGAGGGGAGCTACTTGGGGCTGCCATGGCTCAGTTCCTGACCCCTG 538

QY 4930 GAGCCCTGAACCTGGCTTCAGAGTAGCAAGAGGTTTCTCCAGATGCTGTGAAGGAAATG 4989

DB 539 GAGCCCTGAACCTGGCTTCAGAGTAGCAAGAGGTTTCTCCAGATGCTGTGAAGGAAATG 598

QY 4990 CTTTGTGATAGGAAAAAGGGCGGCTGCTCATTTTATTTTATCTTTTACACTGAATCCC 5049

DB 599 CTTTGTGATAGGAAAAAGGGCGGCTGCTCATTTTATTTTATCTTTTACACTGAATCCC 658

QY 5050 AAAATCATCTTACCACAAAGGGCCAGCTGACTGTTATTTCTTGAGTCAACAAGAGCCAT 5109

DB 659 ANAATCATCTTACCACAAAGGGCCAGCTGACTGTTATTTCTTGAGTCAACAAGAGCCAT 718

QY 5110 GCCATCTCTGGTTTCTCACCTCAGTCATGTCAGAAATTTGTTCAGATCCAGTGGCATCT 5169

DB 719 GCCATCTCTGGTTTCTCACCTCAGTCATGTCAGAAATTTGTTCAGATCCAGTGGCATCT 778

QY 5170 GTGCTCTTGCTGCAACATCTTCTTATTTTCAACTGGCTGCGACATCAAGTGTAACTCTGGC 5229

DB 779 GTGCTCTTGCTGCAACATCTTCTTATTTTCAACTGGCTGCGACATCAAGTGTAACTCTGGC 837

QY 5230 TTCTGGGCCAAG 5241

DB 838 TTCTGGGCCAAG 849

RESULT 4

CA317106

LOCUS

DEFINITION

CA317106

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA317106 817 bp mRNA linear EST 26-NOV-2002
UI-M-FW0-cbm-d-21-0-UI-r1 NIH_BMAP_FW0 Mus musculus cDNA clone
IMAGE: 6811462 5', mRNA sequence.

CA317106

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 817

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6811462"

/tissue_type="whole brain"

/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_FW0"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;

Site 2: Not I; The library was constructed according

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 180 a 254 c 234 g 148 t 1 others
ORIGIN

Query Match 12.8%; Score 792.2; DB 14; Length 817;

Best Local Similarity 98.8%; Pred. No. 6.8e-123;

Matches 808; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

Qy 3009 AGCTCGGCTGGGCTCCAGTGTCCCTGCTGCCATGAGTTGGAGTCCCTCGTGCATCC 3068
Db 1 AGCTCGACTTGGGCCCCCAGTGTCCCTGCTGCCATGAGTTGGAGTCCCTCGTGCATCC 60
Qy 3069 TCGTCCCGAGATTGGTCCCAACCACTCAGATGTGGAAGCAAGGCTGAAGTACACAG 3128
Db 61 TCGTCCCGAGATTGGTCCCAACCACTCAGATGTGGAAGCAAGGCTGAAGTACACAG 120
Qy 3129 CTTATGGGTGGCAGTGTTCAGATTGCGGGGCCACTCAAGAGAAAGATCTCTGGGC 3188
Db 121 CTTATGGGTGGCAGTGTTCAGATTGCGGGGCCACTCAAGAGAAAGATCTCTGGGC 180
Qy 3189 TCAGGCGGGGACCAAACTGGGCGAGCTCTGGGCGAGCTGTGAGCTGCGCCAGGGTAG 3248
Db 181 TCAGGCGGGGACCAAACTGGGCGAGCTCTGGGCGAGCTGTGAGCTGCGCCAGGGTAG 240
Qy 3249 TGGTCCAGGCGGCTCTGACCCGTGTCTGCTGCTCCAGCGGACCGGCGAGACCT 3308
Db 241 TGGTCCAGGCGGCTCTGACCCGTGTCTGCTGCTCCAGCGGACCGGCGAGACCT 300
Qy 3309 GCTGCTCAAGCCCTGTGTATGACGCGATTAAGCAAGCAAGGAGAAAGCGTCCCC 3368
Db 301 GCTGCTCAAGCCCTGTGTATGACGCGATTAAGCAAGCAAGGAGAAAGCGTCCCC 360
Qy 3369 AGCTCGAGGAATCAGTGGAGTGAAGTCAATGTTCATCTCCAGCTTGGTGCATCAA 3428
Db 361 AGCTCGAGGAATCAGTGGAGTGAAGTCAATGTTCATCTCCAGCTTGGTGCATCAA 420
Qy 3429 AGGATGCTGACCTCACTCCAGACCTGGAGCCAGAGAACCACTGACTGCAGAGAC 3488
Db 421 AGGATGCTGACCTCACTCCAGACCTGGAGCCAGAGAACCACTGACTGCAGAGAC 480
Qy 3489 TCTGCTTCCAGCTCTGAGCTGTGGATCTCTCTCAAGGAGCAGACTGGCTGGGACGGGA 3548
Db 481 TCTGCTTCCAGCTCTGAGCTGTGGATCTCTCTCAAGGAGCAGACTGGCTGGGACGGGA 540
Qy 3549 GCTGGGAGGTGCAACCAACCAAGTGGGCGAGAGGCTCACTGCTTGGCAGAGC 3608
Db 541 GCTGGGAGGTGCAACCAACCAAGTGGGCGAGAGGCTCACTGCTTGGCAGAGC 600
Qy 3609 AGCCAGTGCCTCTGCTCTGCTCAGACCTCCAGCCAGCACTCTATAGAGGAGGCC 3668
Db 601 AGCCAGTGCCTCTGCTCTGCTCAGACCTCCAGCCAGCACTCTATAGAGGAGGCC 660
Qy 3669 TGGGAAAGCTGCGCAGGCCCAAGGCCCTGTGCTCTTAACAGTCAAGCCCAAGCTTCCAG 3728
Db 661 TGGGAAAGCTGCGCAGGCCCAAGGCCCTGTGCTCTTAACAGTCAAGCCCAAG-CTTCCAG 719
Qy 3729 GGCCCTCTGCTCTGCTCAGAGTCCCTGAGCAGAGGCGAGATATGCTCAGGAACATG 3788
Db 720 NGCCCTCTGCTCTGCTCAGAGTCCCTGAGCAGAGGCGAGATATGCTCAGGAACATG 779
Qy 3789 CCATGATGGCTACACATGTGTGTAATAGATATCCA 3826
Db 780 CCATGATGGCTACACATGTGTGTAATAGATATCCA 817

RESULT 5

CD348071

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

MUS MUSCULUS (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 748)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabs-rc@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library prepared by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

http://genome.uiowa.edu/distribution/mousefl.html

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

The following repetitive elements were found in this cDNA

sequence: 179-203, >(CA)n#Simple_repeat (matched complement)

Seq primer: pYX-5.

Location/Qualifiers

1..748

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6851477"

/issue_type="whole brain"

/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH-BMAP_FY0"

/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is AGCGACAG. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

BASE COUNT 129 a 202 c 177 g 236 t 4 others

ORIGIN

Query Match 11.9%; Score 737.6; DB 14; Length 748;

Best Local Similarity 98.9%; Pred. No. 1e-113;

Matches 740; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3775 GGCTCAGGACATGCCATGCCTACACATGTGTGTACTAGATATCCATAGTCCT 3834

Db 1 GGCTCAGGACATGCCATGCCTACACATGTGTGTACTAGATATCCATAGTCCT 60

Qy 3835 TCGAGCCTCTTAGGCTCTTTGGCTGGGGTGGGGAGAACTTTACTCTCCCTCATATCT 3894

Db 61 TCGAGCCTCTTAGGCTCTTTGGCTGGGGTGGGGAGAACTTTACTCTCCCTCATATCT 120

Qy 3895 GCATCATACAGAGGGAGCTTGAGACACAGCTCTGTGTAAAGGACACGCTGTAAGTCGT 3954


```

Db      601  CTCTGGAGTCTATGTCTGCGAGGCCAACAGCCCGCCGACCGGTGATTTTGGCCACTGCGGC 660
Qy      975  TGCTGAGCTCCGAGTCTGCTGCTGCCAGCCATCTGCGAGCCACCGAGCGCTCTCGCG 1034
Db      661  TGCTGAGCTCCGAGTCTGCTGCTGCCAGCCATCTGCGAGCCCGCGCGCTCTCGCG 720
Qy      1035  GACGCGGCGCAGCACCGCGCG 1055
Db      721  GACGCGGCGCAGCACCGCGCG 741

RESULT 7
BUB40589
LOCUS
DEFINITION BUB40589 944 bp mRNA linear EST 16-OCT-2002
5', mRNA sequence.
ACCESSION BUB40589
VERSION BUB40589.1 GI:24024989
KEYWORDS Mus musculus (house mouse)
SOURCE EST.
ORGANISM Mus musculus
REFERENCE 1. (bases 1 to 944)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM14039 Row: 0 Column: 04
High quality sequence stop: 535.
FEATURES
source
1..944
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6490299"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 195 a 270 c 227 g 252 t
ORIGIN
Query Match 11.74; Score 720.6; DB 13; Length 944;
Best Local Similarity 96.18; Pred. No. 6.8e-111;
Matches 793; Conservative 0; Mismatches 24; Indels 8; Gaps 5;

Qy      4254  TAAGGTCCAGATGTCGGGTGAACCCAGCTCAGCTCCCTCTCTTTGAGCATCTGTGA 4313
Db      1  TAAGGTCCAGATGTCGGGTGAACCCAGCTCAGCTCCCTCTCTTTGAGCATCTGTGA 60
Qy      4314  CCTAATTTTGAATCTGGGAGTCCCTGGTTGGGAATCTTCTTGGACCCCTGCTCCT 4373
Db      61  CCTAATTTTGAATCTGGGAGTCCCTGGTTGGGAATCTTCTTGGACCCCTGCTCCT 120
Qy      4374  CTCTGCCCTTCTCTCATTTGTTCTGTTGATCTGCTCATGTCATCTTGTCTGATTATCC 4433
Db      121  CTCTGCCCTTCTCTCATTTGTTCTGTTGATCTGCTCATGTCATCTTGTCTGATTATCC 180
Qy      4434  TGGGCGCCCTTCTCTTTCCCATGATGCCCTGATTTCTCTGCTGCTTTCTCTGTC 4493

```

```

Db      181  TGGGGCCCTTCTCTTTTCCCATGATGCCCCCTGATTTCTCCTCACTGCTGTTTTTCAATTTCTGTC 240
Qy      4494  TGGCATGCTTGTCTTTATGTCGTGTTTCTCGTCCCTGAGTTCAACCTATGCAACCCCTTT 4553
Db      241  TGGCATGCTTGTCTTTATGTCGTGTTTCTCGTCCCTGAGTTCAACCTATGCAACCCCTTT 300
Qy      4554  CTTAAACAACATGACTACCTCATGTCGTCTTACAGACCATAGTGTGACCCCTGGTCCCCAC 4613
Db      301  CTTAAACAACATGACTACCTCATGTCGTCTTACAGACCATAGTGTGACCCCTGGTCCCCAC 360
Qy      4614  AGCTCCCTGCGCAACCGCTTCTGGGAGATGAGCCCACTCCAGTAGATCTGGAAGAG 4673
Db      361  AGCTCCCTGCGCAACCGCTTCTGGGAGATGAG-CCCACTCAAGTAGATCTGGAAGAG 419
Qy      4674  ACCCTTGTGGCTTGTCTGCTGCTCCCTCCCTTGGTGTTCAGATGAGAGGTTTCTATGG 4733
Db      420  A-CCCTTGTGGCTTGTCTGCTGCTCCCTCCCTTGGTGTTCAGATGAGAGGTTTCTATGG 478
Qy      4734  AAGAGATGATCGAGGCTGCA CAGGGGAACCCCAAGAGAGGAGTGGAGTGAACCAAG 4793
Db      479  AAGAGATGATCGAGGCTGCA CAGGGGAACCCCAAGAGAGGAGTGGAGTGAACCAAG 538
Qy      4794  AGGCTGAAAAAATGGCTGCCACCATCTGCACAGAGAGATGGGTGTGCTTTTGAGC 4853
Db      539  AGGCTGAAAAAATGGCTGCCACCATCTGCACAGAGAGATGGGTGTGCTTTTGAGC 598
Qy      4854  TGCAGTCTCTGCTGAAAATGAAAGGGGTGAGGAGAGGAGGAGTACTGGGGCTGCATGGCT 4913
Db      599  TGCAGTCTCTGCTGAAAATGAAAGGGGTGAGGAGAGGAGGAGTACTGGGGCTGCATGGCT 658
Qy      4914  CAGTTCCTCGACCCCTGGAGCCCTGAAACCTGGCTCTCAGATGAGCAAGAGATTTCTCTCAAG 4973
Db      659  CAGTTCCTCGACCCCTGGAGCCCTGAAACCTGGCTCTCAGATGAGCAAGAGATTTCTCTCAAG- 717
Qy      4974  ATGCTGTAGGAGAGTCTTTGCAATAGGAGAGAGGAGGAGGAGTGGCTCATTTTA-TTTTATCT 5031
Db      718  ATGCTGTAGGAGAGGAGCCTTTGCAATAGGAGAGAGGAGGAGTGGCTCATTTTAATTTTAACCT 777
Qy      5032  TTCTTTTACACTGAA---TCCCAAAATCATCTTTACCAAAAGGGGCC 5073
Db      778  TGCCTTACACCGGATCCCAAAATCATCTTTACCAAAAGGGGCC 822

BUB405518 808 bp mRNA linear EST 09-OCT-2002
UI-M-FOO-cac-e-03-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
IMAGE: 6409322 5', mRNA sequence.
BUB405518
BUB405518.1 GI:23635078
EST.
Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 808)
AUTHORS NIH-MGC http://imgc.ncbi.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
cDNA Library Preparation: Dr. Jim Lin, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
The following repetitive elements were found in this cDNA
sequence: 30-59, >(TAAAA)n#simple_repeat (matched complement)
Seq primer: pyX-5.

```



```

QY 2499 TGGCTCCGTGTAGAAAGCTCCACCCCTGCGTGAACGGGCTTCAACACCTCTTCTGACCT 2558
Db 540 TGGCTCCGTGTAGAAAGCTCCACCCCTGCGTGAACGGGCTTCAACACCTCTTCTGACCT 599
QY 2559 GCGCTGAGCCCTGTGACACCATCCACCGTTGGTTACACTGGTGTCCCGCCGACGGGCC 2618
Db 600 GCGCTGAGCCCTGTGACACCATCCACCGTTGGTTACACTGGTGT-CCCCCGAGGACC 658
QY 2619 CAATGGTGTAGATTGGAGTATCTAAATTTCTTACAGCAACACACCCAGCCCGGAACA 2678
Db 659 CAATGGTGTAGATTGGAGTATCTAAATTTCTTACAGCAACAA-CACACCCAGCCCGGAACA 717
QY 2679 CCAGTGGACATGTCTCACACAG 2701
Db 718 CCAGTGGACATGTCTCACACAG 740

RESULT 11
CB524810
LOCUS
DEFINITION
  UI-M-FY0-cey-g-12-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
  IMAGE: 6843781 5', mRNA sequence.
ACCESSION
  CB524810
VERSION
  CB524810.1 GI:29358165
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1. (bases 1 to 713)
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: Dr. Jim Lin, University of Iowa
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Distribution information can be found at
  http://genome.uiowa.edu/distribution/mousefi.html
  This clone was contributed by the Brain Molecular Anatomy Project
  (BMAP)
  Seq primer: pYX-5.
FEATURES
  Location/Qualifiers
  1..713
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6"
  /db_xref="taxon:10090"
  /clone="N1H BMAP_FY0"
  /tissue_type="whole brain"
  /dev_stages="embryo 13.5,14.5,16.5,17.5dpc"
  /lab_host="DH10B (T1 phage resistant)"
  /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
  Site 2: Not I; The library was constructed according
  Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
  1996. Denatured RNA was size fractionated on a 1% agarose
  gel. First strand cDNA synthesis was primed with oligo-dT
  primer containing a Not I site. Double strand cDNA was
  size selected according to mRNA size fraction, ligated
  with EcoR I adaptor, digested with NotI and then cloned
  directionally into pYX-Asc vector. The library tag
  sequence located between the Not I site and the polyA tail
  is AGCAGACAG. This library was created for the University
  Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
  Developing Mouse Nervous System', supported by National
  Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
  program coordinator."
  141 a 235 c 185 g 148 t 4 others
  BASE COUNT
  ORIGIN

```

```

Query Match 11.2%; Score 690; DB 14; Length 713;
Best Local Similarity 98.6%; Pred. No. 9.7e-106;
Matches 704; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 2489 ATGGGCGCTTTGGCTCGTGTAGAACGCTCCACCCCTGCGTGAACGGGCTTCAACACCTC 2548
Db 1 ATGGGCGCTTTGGCTCGTGTAGAACGCTCCACCCCTGCGTGAACGGGCTTCAACACCTC 59
QY 2549 CTTCTGACCTGCGCTGACGCCCCCTGACACCATCCACCGTTGGTTACACTGGTGTCCCG 2608
Db 60 CTTCTGACCTGCGCTGACGCCCCCTGACACCATCCACCGTTGGTTACACTGGTGTCCCG 119
QY 2609 CCACGAGCCCAATGGTGTAGATTGTGGAGTATCTAAATTTCTTACAGCAACACACACCC 2668
Db 120 CCACGAGCCCAATGGTGTAGATTGTGGAGTATCTAAATTTCTTACAGCAACACACACCC 179
QY 2669 AGCCCGAACACCACTGACACATGCTCACCACAGAGGGAACAATCTTCAAGTGCAGAGTCC 2728
Db 180 AGCCCGAACACCACTGACACATGCTCACCACAGAGGGAACAATCTTCAAGTGCAGAGTCC 239
QY 2729 ATGGCCTTAGAGAGTACACTCGGTATTTCTTCAAGATGGGAGCCCGCACAGAGTGGGGC 2788
Db 240 ATGGCCTTAGAGAGTACACTCGGTATTTCTTCAAGATGGGAGCCCGCACAGAGTGGGGC 299
QY 2789 CTGGGCGCTTTTCCCGCTTGCAGGATGTGATTCTCTGCAAGAGACATTTCTCAGACTCCT 2848
Db 300 CTGGGCGCTTTTCCCGCTTGCAGGATGTGATTCTCTGCAAGAGACATTTCTCAGACTCCT 359
QY 2849 TGGATGTGACCGCGCTCAACGGGCAATCATGTGGTGTCTGCTGGGCTTCTCTGCTCCTCC 2908
Db 360 TGGATGTGACCGCGCTCAACGGGCAATCATGTGGTGTCTGCTGGGCTTCTCTGCTCCTCC 419
QY 2909 TGGCTTCGATGTGTGCTGCGCTACGACAAAGCTCCACAGGAGAGCCCTTCCCGGATTGT 2968
Db 420 TGGCTTCGATGTGTGCTGCGCTACGACAAAGCTCCACAGGAGAGCCCTTCCCGGATTGT 479
QY 2969 CCTCTCAGGACACCCAGGAAACCCAGCGCTCTACACAGAGCTCGGCTTGGGCTCCCA 3028
Db 480 CCTCTCAGGACACCCAGGAAACCCAGCGCTCTACACAGAGCTCGGCTTGGGCTCCCA 539
QY 3029 GTGTCCCTGTGCCATAGTGTGAGTCCCTCTGTGCAATCTCTGCTCCCGAGGATTGTGCC 3088
Db 540 GTGTCCCTGTGCCATAGTGTGAGTCCCTCTGTGCAATCTCTGCTCCCGAGGATTGTGCC 599
QY 3089 CACCACCTCAGATGTGGAAGACAGGCTGAGTACAGAGCTTATGCGTGCAGTGTCTT 3148
Db 600 CACCACCTCAGATGTGGAAGACAGGCTGAGTACAGAGCTTATGCGTGCAGTGTCTT 659
QY 3149 CAGATTGCGGGGGCCACTCCAGAGAGAAAGATCTCTGCGGCTCAGCGAGGGGAC 3202
Db 660 CAGATTGCGGGGGCCACTCCAGAGAGAAAGATCTCTGCGGCTCAGCGAGGGGAC 713

RESULT 12
LOCUS
DEFINITION
  UI-M-FY0-cft-p-13-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
  IMAGE: 6852446 5', mRNA sequence.
ACCESSION
  CB511162
VERSION
  CB511162.1 GI:31142737
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1. (bases 1 to 699)
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: Dr. Jim Lin, University of Iowa

```

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefi.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. .699
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6852446"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP FYO"

/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGGAGAGAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

BASE COUNT 135 a 221 c 202 g 139 t 2 others

ORIGIN

Query Match 11.1% Score 685.8; DB 14; Length 699;

Best Local Similarity 98.7% Pred. No. 4.9e-105;

Matches 690; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

2753 ATTCTTTCAAGATGGGAGCCGACACAGAGTGGGGCTGGGCTTTTCCCGCTTGCGG 2812

1 ATTCTTTCAAGATGGGAGCCGACACAGAGTGGGGCTGGGCTTTTCCCGCTTGCGG 60

2813 ATGTGATTAATCTGCAAGAGACATTTCTCAGACTCTTTGGATGTCACGCCGTCAAGGCA 2872

61 ATGTGATTAATCTGCAAGAGACATTTCTCAGACTCTTTGGATGTCACGCCGTCAAGGCA 120

2873 TCATCGTGGGTGTCGCTGGGCTTTCTGCTCTCTGGCTTGGCTGATGTGCTGGCTAC 2932

121 TCATCGTGGGTGTCGCTGGGCTTTCTGCTCTCTGGCTTGGCTGATGTGCTGGCTAC 180

2933 GACAAAGCTCCACAGAGAGCCCTTCCGGATGTCCTCTCAGGACCCCGAGGAACC 2992

181 GACAAAGCTCCACAGAGAGCCCTTCCGGATGTCCTCTCAGGACCCCGAGGAACC 240

2993 CAGCGCTTACACAGAGCTCGGCTTGGGCTCCAGTGTCCCTGTGCTCCCATGAGTTGG 3052

241 CAGCGCTTACACAGAGCTCGAATGGGCCCCCAGTGTCCCTGTGCTCCCATGAGTTGG 300

3053 AGTCCCTCGTGCATCTCTGCTCCCGAGTTGGTCCCGAGGCTCAGATGTGGAAGA 3112

301 AGTCCCTCGTGCATCTCTGCTCCCGAGTTGGTCCCGAGGCTCAGATGTGGAAGA 360

3113 AGGTGAAGTACACAGAGCTTATGGGTGGCAGTGTTCAGATTTGGCGGGCCACTTCCAGA 3172

361 AGGTGAAGTACACAGAGCTTATGGGTGGCAGTGTTCAGATTTGGCGGGCCACTTCCAGA 420

3173 GAAAGATCTCTGGGCTCAGGAGGGGACCAACTGGGAGGCTCTCTGGGAGGCTGTG 3232

421 GAAAGATCTCTGGGCTCAGGAGGGGACCAACTGGGAGGCTCTCTGGGAGGCTGTG 480

3233 AGCTGCCCGAGGTTAGTGTCAAGCGCGGCTCTGACCGGCTCTGCTGCTCCAGCGG 3292

481 AGCTGCCCGAGGTTAGTGTCAAGCGCGGCTCTGACCGGCTCTGCTGCTCCAGCGG 540

3293 GAACCGGGCAGACACTGCTGCTGCAAGCCCTGCTGTATGACGGCATAAAGAGCAACGGA 3352

541 GAACCGGGCAGACACTGCTGCTGCAAGCCCTGCTGTATGACGGCATAAAGAGCAACGGA 600

3353 GAAAGAGCGGTCGCCCGAGCTGCAGGAATCAGGTGGAAGTCAATGTCCTCCACTCCG 3412

601 GAAAGAGCGGTCGCCCGAGCTGCAGGAATCAGGTGGAAGTCAATGTCCTCCACTCCG 660

3413 ACTTCGGTGCATCCCAAGGATGTCCTGACCTCCACTCC 3451

661 ACTTCGGTGCATCCCAAGGATGTCCTGACCTCCACTCC 699

RESULT 13

BI185886

LOCUS

DEFINITION

BI185886

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Pomp, D

Department of Animal Science

University of Nebraska, Lincoln

Lincoln, NE 68583-0908, USA

Tel: 402 472 6416

Fax: 402 472 6362

Email: dpomp@unl.edu

Oligo-dT track not found, Not I site shown in beginning of sequence

is likely internal to the message.

Seq primer: M13 -29

POLYA=No.

Location/Qualifiers

1. .900

/organism="Sus scrofa"

/mol_type="mRNA"

/strain="University of Nebraska, Lincoln Swine Selection

Lines"

/db_xref="taxon:9823"

/clone="UNL-P-FN-df-h-02-0-UNL"

/dev_stage="ADULT"

/lab_host="DH10B (Life Technologies)"

/clone_lib="UNL-P-FN"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The UNL-P-FN

library is a normalized library representing porcine

ovarian follicles, ranging between 2.0 to 10.0 mm in

diameter, collected during 7 days of the follicular phase

of the pig estrous cycle. This library was derived from

the library UNL-P-F2. The tag is a string of 5-6

nucleotides present between the Not I site and the

oligo-dT track. The library was constructed as described

by Bonaldo, Lennon and Soares, Genome Research 6: 791-806

, 1996.

BASE COUNT 149 a 304 c 291 g 150 t 6 others

ORIGIN

Query Match 10.4%; Score 639.6; DB 12; Length 900;
Best Local Similarity 86.1%; Pred. No. 2.7e-97;
Matches 719; Conservative 0; Mismatches 115; Indels 1; Gaps 1;

QY 739 GCAGCCCGAGAGACACACAGGTAGTGTCTGACAGATGTAGTATGAGTGTGCTGCC 798
Db 7 GCTGCCCCCGAGAACACACAGTGTGTGTCGCGCCAGAGCGTGTATGAGTGTGCTGCC 66
QY 799 TCTGCTGACCCACACCCCTTTTGTCTGCTGGTCCGACAGGATGGAAGCCTATCTCCACG 858
Db 67 TCGGCTGACCCACACCCCTATCTGTCTGGTCCGACAGGAGCGGAGCCTATTTCTACC 126
QY 859 GATGTATCTGTTCTGGGCGGACCAATCTATCTATCTGTCGACGCGCGAGCTCTGCACT 918
Db 127 GATGTATCTGTTCTGGGCGGACCAATCTATCTATCTGTCATCAGCGCGCANCC-CGGCACT 185
QY 919 GGAGTCTATGTCGCGAGCCCAACAGCCCTCTCAGCGTGTACTTCCCACTGCGGCTCT 978
Db 186 GCGGTCTAGTCTGAGAGCCCAACAGCCCTCTCAGCGTGTACTTCCCACTGCGGCTCT 245
QY 979 GAGCTCCGAGTCTTGTCTGCCCCAGCCATCTCCGAGGACCCGAGCGCTCTCCGCGGAG 1038
Db 246 GAGCTCCGCGTCTGCGGCTCTGCGCATCTGCGAGGCGCCGAGCGCTATCTCGAGCG 305
QY 1039 CGGCGCAGCACCGCGCTCTGTGTGCGGCGCTCCGCGGAGCCACGCGCGCGCTGCGAC 1098
Db 306 AGGGCGAGCACCGCGCTCTGTGTGCGCGCAACGCGGTGAGCGCGCGCGCTGCGG 365
QY 1099 TGGCTGACAGCGGATCCGTTGCGACCAATGCGGCGCTCAAGTGCAGGCGCGTGC 1158
Db 366 TGGCTGCAACCGGGCGCGCTGCGGCGCCACGCGCGCTCAAGTGCAGGCGCGCGCG 425
QY 1159 GCGAGTCTGTCATCACTCAGATCGGCTGCGAGGAGCTGCGTCTACAGTGCCTAGCA 1218
Db 426 AGCAGCTGCTGTCATCAGCAGATCGGCTGCGAGGATGCGGCTACTACAGTGCCTG 485
QY 1219 GAAACAGCGCGGAACTGCTGTGCGCTGCGGCGCTGCGGCTGAGTGTGCGGAGGG 1278
Db 486 GAGAACAGCGCGGACGCGCTGCGCGCGCGCTGCGGCTGAGTGTGCGGAGGGT 545
QY 1279 CTGCCCCAGGCGCGGCTGCGGTCAAGCAGCGCGCTGAGAGCTCTCTGTGTGCTGT 1338
Db 546 CTGCCCCAGGCGCGGCTCAAGCAGCGCGCTGAGAGCTCTCTGTGTGCTGTGT 605
QY 1339 GCCTGGAGCGGCTGAGTGTGACAGCGAGCAATCATTTGGCTCTCTCTTCACTACCA 1398
Db 606 GCCTGGAGCGGCGGCTGACAGCGAGCAATCATTTGGCTCTCTCTTCACTACCA 665
QY 1399 AAGCAAGGGGAGTGGACAAATGTGAGTACCAGTTTGCAGTAAACAAATGACACACAG 1458
Db 666 AAGCAAGGGGAGTGGACAAATGTGAGTACCAGTTTGCAGTAAACAAATGACACACAG 725
QY 1459 CTGAGGTTTGGGACCTGGAACCCAAACGGAATATGATTTCTAGTGTGCGCTACTCC 1518
Db 726 CTACAGGTTTGGGACCTGGAACCCAAACGGAATATGATTTCTAGTGTGCGCTACTCA 785
QY 1519 CAGCTGGGGGCGAGCGGACCTCCAGCGCGCTGCTGCTATACACTGAGCGATG 1573
Db 786 CAGCTGGGGGCGAGCGGACCTCCAGCGCGCTGCTGCTATACACTGAGCGATG 840

RESULT 14
CB518504
LOCUS
DEFINITION
IMAGE: 6837653 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CB518504 631 bp mRNA linear EST 28-MAR-2003
UI-M-GH0-ccc-h-04-0-UI-r1 NIH_BMAP_GH0 Mus musculus cDNA clone
IMAGE: 6837653 5', mRNA sequence.
CB518504
CB518504.1 GI:29351859
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 631)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouseefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES
source

Location/Qualifiers

1..631
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6837653"
/tissue_type="Whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP GH0"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAATGTAAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

BASE COUNT 151 a 183 c 180 g 117 t

Query Match 10.2%; Score 629.4; DB 14; Length 631;
Best Local Similarity 99.8%; Pred. No. 1.5e-95;
Matches 630; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1406 GGGGAGTGGACAATGTGGAGTACCAGTTTGCAGTAAACAAATGACACACAGAGCTGCAGG 1465
Db 1 GGGGAGTGGACAATGTGGAGTACCAGTTTGCAGTAAACAAATGACACACAGAGCTGCAGG 60
QY 1466 TTCGGGACCTTGAACCCCAACACGGAATATGATTTCTAGTGTGCGCTTACTCCAGCTGG 1525
Db 61 TTCGGGACCTTGAACCCCAACACGGAATATGATTTCTAGTGTGCGCTTACTCCAGCTGG 120
QY 1526 GGGGACGCGGAACCTCCAGCGCGCGCTGGTGCATACACTGACGATGTCCCGCGGAG 1585
Db 121 GGGGACGCGGAACCTCCAGCGCGCGCTGGTGCATACACTGACGATGTCCCGCGGAG 180
QY 1586 CACCCGAGCTTACCTTGTTCAGCGCCCAACCCCTCGGACATCAGGCTGCGATGGCTGCC 1645
Db 181 CACCCGAGCTTACCTTGTTCAGCGCCCAACCCCTCGGACATCAGGCTGCGATGGCTGCC 240
QY 1646 TGCCCTCCAGCTGAGCAATGGAAGGCTGCTGGAATGACACACAACTTACGTTAACT 1765
Db 241 TGCCCTCCAGCTGAGCAATGGAAGGCTGCTGGAATGACACACAACTTACGTTAACT 300
QY 1706 AGGAGATCAGGTTTCTCCACCGGAGTGTCTGGAATGACACACAACTTACGTTAACT 1765
Db 301 AGGAGATCAGGTTTCTCCACCGGAGTGTCTGGAATGACACACAACTTACGTTAACT 360
QY 1766 CACTTCAGCCAAACAAAGTGTACCGAGTCCGGATTTTCAGCTGCGCAGCTGGCTGCTATG 1825

Qy 5413 ACACATAAGATTATTAATGCTGGCTTCGTATCTCATTAAGACAGGATTGTCACTTGAAC 5472
Db |||||
181 ACACATAAGATTATTAATGCTGGCTTCGTATCTCATTAAGACAGGATTGTCACTTGAAC 240
Qy 5473 TACTTCTATAGATTCAAGTGGCCACGCGCAACACCAACCGTATGTTTCTTCATTGCTCT 5532
Db |||||
241 TACTTCTATAGATTCAAGTGGCCACGCGCAACACCAACCGTATGTTTCTTCATTGCTCT 300
Qy 5533 GAAGGTCAAGAGCCTCATTTTGTCTTCTGGTTAGATTCTTCTCCTTGCCTTGAATG 5592
Db |||||
301 GAAGGTCAAGAGCCTCATTTTGTCTTCTGGTTAGATTCTTCTCCTTGCCTTGAATG 360
Qy 5593 AAATAACCGTTTTAAACAGTAGGCTCTTAGCATCACACATAGTCAATTCCTCATGTTCT 5652
Db |||||
361 AAATAACCGTTTTAAACAGTAGGCTCTTAGCATCACACATAGTCAATTCCTCATGTTCT 420
Qy 5653 TGTTTAAACAGCATTTGAGGTTCTGGGTTTAAATTAATAGCTGCAAAATGAGACAAATTA 5712
Db |||||
421 TGTTTAAACAGCACT--GAGGTTCTGGTTTAAATTAATAGCTGCAAAATGAGACAAATTA 478
Qy 5713 TAACCCATTAGCTGGTGGAAATTTG--TTCTCAAAAGCAAAATAGTAATAAATCTGGTA 5771
Db |||||
479 TAACCCATTAGCTGGTGGAAATTTGTTCTCAAAAGCAAAATAGTAATAAATCTGGTA 538
Qy 5772 TCTGCTATAACTCACAGTTGATAGAAAAGTAGCCAGAACTCACATAGCATTATATATGAT 5831
Db |||||
539 TCTGCTATAACTCACAGTTGATAGAAAATAGCCAGAACTCACATAGCATTATATATGAT 598
Qy 5832 TGGGGTTCTGAGTAACTGGGAGTGTAGCTTTGTGACTTTGTAGCACAGGCTTTATTA 5891
Db |||||
599 TTGGGTTCTGAGTAACTGGGAGTGTAGCTTTGTGAAATTTGGAACAAACAGGCTCTTATTA 658
Qy 5892 GGAAGTCTGTTG 5904
Db |||||
659 AGAAAAGTTGGTG 671

Search completed: October 15, 2003, 11:38:58
Job time : 14070.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 06:22:55 ; Search time 1163.58 Seconds
(without alignments)
14327.989 Million cell updates/sec

Title: US-09-754-997A-1
Perfect score: 6176
Sequence: 1 atggcgggcggaacggg.....taataaactatttccagac 6176

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues 5105512
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_13Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6176	100.0	6176	22	Mouse Nope (neighb
2	2804.2	45.4	3753	22	Nucleotide sequenc
3	2796	45.3	2796	22	Mouse Nope (neighb
4	2681.8	43.4	3741	24	DNA encoding human
5	2118.8	34.3	3450	24	Human polynucleoti
6	1522.4	24.7	2223	22	Nucleotide sequenc
7	825	13.4	825	22	Mouse Nope (neighb
8	337.2	5.5	680	21	Human secreted pro

9	316.6	5.1	2976	25	ABX71437	Human cDNA encodin
10	316.6	5.1	3210	25	ABX71432	Human cDNA encodin
11	316.6	5.1	3219	25	ABX71440	Human cDNA encodin
12	316.6	5.1	3874	25	ABX71447	Human cDNA encodin
13	316.4	5.1	2700	24	AAD36777	Human immunoglobul
14	315.4	5.1	3453	25	ABX71435	Human cDNA encodin
15	289	4.7	2695	23	AAS76339	DNA encoding novel
16	288	4.7	288	22	AAD10030	Mouse Nope (neighb
17	252	4.1	252	22	AAD10032	Mouse Nope (neighb
18	249	4.0	249	22	AAD10028	Mouse Nope (neighb
19	249	4.0	249	22	AAD10029	Mouse Nope (neighb
20	246	4.0	246	22	AAD10031	Mouse Nope (neighb
21	243	3.9	243	22	AAD10024	Mouse Nope (neighb
22	217	3.5	2481	25	ABX71439	Human cDNA encodin
23	217	3.5	2715	25	ABX71434	Human cDNA encodin
24	217	3.5	2724	25	ABX71441	Human cDNA encodin
25	217	3.5	2958	25	ABX71436	Human cDNA encodin
26	213.8	3.5	2139	25	ABX71442	Human cDNA encodin
27	213.8	3.5	2382	25	ABX71445	Human cDNA encodin
28	201.8	3.3	563	24	ABQ25578	Oligonucleotide fo
29	201.8	3.3	563	24	ABQ25579	Oligonucleotide fo
30	200.2	3.2	3330	23	AAS76658	DNA encoding novel
31	195	3.2	195	22	AAD10027	Mouse Nope (neighb
32	192	3.1	192	22	AAD10025	Mouse Nope (neighb
33	190.4	3.1	563	24	ABQ25580	Oligonucleotide fo
34	190.4	3.1	563	24	ABQ25581	Oligonucleotide fo
35	189	3.1	189	22	AAD10026	Mouse Nope (neighb
36	151.2	2.4	1875	25	ABX71443	Human cDNA encodin
37	119	1.9	909	25	ABX71438	Human cDNA encodin
38	119	1.9	1143	25	ABX71433	Human cDNA encodin
39	117	1.9	1644	25	ABX71444	Human cDNA encodin
40	117	1.9	1887	25	ABX71446	Human cDNA encodin
41	99.4	1.6	2450	23	AAS76657	DNA encoding novel
42	93.8	1.5	2560	25	ABX71188	Novel human cDNA s
43	85.6	1.4	439	24	ABS69793	Novel murine polyn
44	75	1.2	370	22	AAS33787	Human cDNA encodin
45	71	1.1	4277	21	AAA37039	Human PRO1190 (UNQ

ALIGNMENTS

RESULT 1	
AAD10021	
ID	AAD10021 standard; cDNA; 6176 BP.
XX	
AC	AAD10021;
XX	
DT	12-SEP-2001 (first entry)
XX	
DE	Mouse Nope (neighbour of punc ell) cDNA.
XX	
KW	Mouse; Nope; neighbour of punc ell; cytostatic; neuroprotective; vaccine;
KW	gene therapy; cerebroprotective; colonic cancer; mental retardation;
KW	tumour suppressor; chromosome 9; transgenic animal; genetic disorder;
KW	obesity; Bardet-Biedl syndrome; autosomal recessive disorder; retinitis;
KW	polydactyly; pigmentosa; hypogonadism; hypertension; diabetes mellitus;
KW	renal anomaly; cardiovascular anomaly; ss.
OS	Mus musculus.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..3759
FT	/*tag= a
FT	/product= "Mouse Nope protein"
FT	1..63
FT	/*tag= b
FT	64..3756
FT	/*tag= c
FT	/product= "Mature mouse Nope protein"
FT	1..3756
FT	/*tag= d
FT	/note= "This region is specifically referred as

FT FT SEQ ID NO: 45 in the specification"
 FT 64..2859
 FT /tag= e
 FT /note= "Extracellular domain"
 FT 139..381
 FT /tag= f
 FT /label= Ig1
 FT /note= "Immunoglobulin domain 1"
 FT 463..654
 FT /tag= g
 FT /label= Ig2
 FT /note= "Immunoglobulin domain 2"
 FT 766..954
 FT /tag= h
 FT /label= Ig3
 FT /note= "Immunoglobulin domain 3"
 FT 1039..1233
 FT /tag= i
 FT /label= Ig4
 FT /note= "Immunoglobulin domain 4"
 FT 1285..1533
 FT /tag= j
 FT /label= Fn1
 FT /note= "Fibronectin-type III domain 1"
 FT 1579..1827
 FT /tag= k
 FT /label= Fn2
 FT /note= "Fibronectin-type III domain 2"
 FT 1888..2175
 FT /tag= l
 FT /label= Fn3
 FT /note= "Fibronectin-type III domain 3"
 FT 2248..2493
 FT /tag= m
 FT /label= Fn4
 FT /note= "Fibronectin-type III domain 4"
 FT 2542..2793
 FT /tag= n
 FT /label= Fn5
 FT /note= "Fibronectin-type III domain 5"
 FT 2932..3756
 FT /tag= o
 FT /note= "Intracellular domain"
 FT 3819..5682
 FT /tag= p
 FT 3'UTR
 FT WO200149714-A2.
 FT 12-JUL-2001.
 FT 26-OCT-2000; 2000WO-US29698.
 FT 04-JAN-2000; 2000US-0174496.
 FT 19-MAY-2000; 2000US-0205789.
 FT (NEUR-) NEUROSCIENCES RES FOUND INC.
 FT Salbaum JM;
 FT WPI; 2001-441846/47.
 FT P-PSDB; AAE05251.
 FT Murine Nope polypeptides and nucleic acids useful for preventing,
 FT diagnosing and treating colonic cancer and Bardet-Biedl syndrome -
 FT Claim 10; Fig 2A; 99pp; English.

CC The present invention relates to Nope (neighbour of punc ell) which is
 CC used in the prevention, treatment and diagnosis of diseases associated
 CC with inappropriate Nope expression such as cancers especially colonic
 CC cancer and genetic disorders, as Nope is thought to be a tumour
 CC suppressor. Nope gene is located on chromosome 9 and is used in gene
 CC therapy. Nope is used as vaccine. Nope gene may be administered to treat

CC diseases by rectifying mutations or deletions in a patient's genome that
 CC affect the activity of Nope by expressing inactive proteins or to
 CC supplement the patients own production of Nope polypeptides. Nope gene
 CC is used to study the expression and function of Nope polypeptides and
 CC their role in metabolism through the creation of transgenic animal
 CC models. The anti- Nope antibodies and Nope antagonists may also be used
 CC to down regulate Nope expression and activity for the treatment of
 CC Bardet-Biedl syndrome which is an autosomal recessive disorder
 CC characterised by mental retardation, obesity, polydactyly, retinitis
 CC pigmentosa and hypogonadism. Patients with Bardet-Biedl syndrome have a
 CC high incidence of hypertension, diabetes mellitus and renal and
 CC cardiovascular anomalies. The present sequence is mouse Nope (neighbour
 CC of punc ell) cDNA.

XX Sequence 6176 BP; 1316 A; 1756 C; 1634 G; 1470 T; 0 other;

Query Match 100.0%; Score 6176; DB 22; Length 6176;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 6176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCGCGGGCGGACACGGGGCGGGCTCTGGTGTGACCTTCTGCTGCTGTCCGGG 60
 DB |||||
 QY 1 ATGGCGCGGGCGGACACGGGGCGGGCTCTGGTGTGACCTTCTGCTGCTGTCCGGG 60
 DB |||||
 QY 61 CGCGGGAGCTGCCATTGCCCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGACCC 120
 DB |||||
 QY 61 CGCGGGAGCTGCCATTGCCCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGACCC 120
 DB |||||
 QY 121 CTGCAAGTATCTCTGGGCGCTGAGCAGGCTGTGGTGTGACCTGTGACCTTTGGGGGTACA 180
 DB |||||
 QY 121 CTGCAAGTATCTCTGGGCGCTGAGCAGGCTGTGGTGTGACCTGTGACCTTTGGGGGTACA 180
 DB |||||
 QY 181 GCTGTGGGCTCTCCGACAGGCTGACATGGACCAAGGATGGAGACACTGTACTAGAGCAT 240
 DB |||||
 QY 181 GCTGTGGGCTCTCCGACAGGCTGACATGGACCAAGGATGGAGACACTGTACTAGAGCAT 240
 DB |||||
 QY 241 GAGAACTGCACTGTCTACCCCAATGGCTCCCTGTGGCTGTCTCTCAACCTTAGAGCAAGAA 300
 DB |||||
 QY 241 GAGAACTGCACTGTCTACCCCAATGGCTCCCTGTGGCTGTCTCTCAACCTTAGAGCAAGAA 300
 DB |||||
 QY 301 GACAGGATGATGAGGAGCTCTTAGGATCTGGAGGTGCTACTAGGGGAGCTATTCTCTGT 360
 DB |||||
 QY 301 GACAGGATGATGAGGAGCTCTTAGGATCTGGAGGTGCTACTAGGGGAGCTATTCTCTGT 360
 DB |||||
 QY 361 CTGGGCCACAGCCCGCTAGGAGTGTGGCCAGCAGTTGTGTCTCAAGCTTGCACCA 420
 DB |||||
 QY 361 CTGGGCCACAGCCCGCTAGGAGTGTGGCCAGCAGTTGTGTCTCAAGCTTGCACCA 420
 DB |||||
 QY 421 CTCGAAGACTTCTCTTGCAACCCCGAGTCCAGATTTGTGGAGGAGAAACGGGACAGCACGC 480
 DB |||||
 QY 421 CTCGAAGACTTCTCTTGCAACCCCGAGTCCAGATTTGTGGAGGAGAAACGGGACAGCACGC 480
 DB |||||
 QY 481 TTTGAATGCCACACCAAGGGCTTTCAGCCCCCATCATTTACTTGGGAAAGAGCAAGGTG 540
 DB |||||
 QY 481 TTTGAATGCCACACCAAGGGCTTTCAGCCCCCATCATTTACTTGGGAAAGAGCAAGGTG 540
 DB |||||
 QY 541 ACCGTGCTGAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCTCTAGAT 600
 DB |||||
 QY 541 ACCGTGCTGAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCTCTAGAT 600
 DB |||||
 QY 601 GTCCAGGACAGTGATGAGGCTCTTACCGCTCGGTGGCCCAACCAATTCAGCCCGCCACCA 660
 DB |||||
 QY 601 GTCCAGGACAGTGATGAGGCTCTTACCGCTCGGTGGCCCAACCAATTCAGCCCGCCACCA 660
 DB |||||
 QY 661 TTTAGCCAGGAGGCTCTGCTCACTGTGGCCCTCAGAGGGTCTTTTGAGGCTTACCAAGGGG 720
 DB |||||
 QY 661 TTTAGCCAGGAGGCTCTGCTCACTGTGGCCCTCAGAGGGTCTTTTGAGGCTTACCAAGGGG 720
 DB |||||
 QY 721 CAGGATGTGCTCATTTGTGGCAGCCCGGAGACACCAAGTGTCTCTGACAGATGTA 780
 DB |||||
 QY 721 CAGGATGTGCTCATTTGTGGCAGCCCGGAGACACCAAGTGTGTCTGACAGATGTA 780
 DB |||||
 QY 781 GTGATGGAGTGGTGGGCTCTGCTGACCCACCCCTTTTGTGTCTCTGGGTCCGACAGGAT 840

QY	3001	TACAAAGAGCTGGCTTGGCCCTCCAGTGTCCCTGCTGCCCATGAGTTGGAGTCCCTC	3060
DB	3001		
QY	3061	GTGCATCTCTGCTCCCAAGATTTGGTTCCTCCACACCTCTCAGATGTGGAAGCAAGGGCTGAA	3120
DB	3061		
QY	3121	GTACACAGCCCTTATGGTGTGAGTGTTCAGATTTGGCGGGGCCACATCCCAAGAGAAAGATC	3180
DB	3121		
QY	3181	TCCTGGGCTCAGGCAGGGGACCAAACTGGGCGAGGCTCTGGGCGAGCTGTGAGCTGCC	3240
DB	3181		
QY	3241	CAGGCTAGTGGTCCAAAGGCGGGCTCTGACCCGCTGCTCTGCTGCCCTCAGCGGGAACCGGG	3300
DB	3241		
QY	3301	CAGACACTGTGCTGCAAGCCCTGGTGTATGACGGCATAAAGAGCAACGGGAGAAAGAG	3360
DB	3301		
QY	3361	CCGTCCCCAGCCTGCAAGCCCTGGTGTATGACGGCATAAAGAGCAACGGGAGAAAGAG	3420
DB	3361		
QY	3421	GCATCCAAAGGATGTCTGACCTCCACCTCCAAAGCCTGGAGCCAGAGGAACCACTGACT	3480
DB	3421		
QY	3481	GCAGAGACTCTGCTTCCAGTCTGGAGCTGTGATCTGTCTCAAGGAGCAGACTGGCTG	3540
DB	3481		
QY	3541	GGCAGGGAGCTGGGAGGTTGCCAAACCAACAGTGGGCGCAGAGGGCTCACCTGCTTG	3600
DB	3541		
QY	3601	CCAGAAAGCAGTGGCTCTGCTCTGCTCAGAGCTCCAGGCCAGCAGCTGTATAGAG	3660
DB	3601		
QY	3661	GAGGCCCTTGGGAAAAGCTGCCAGCCCAAGGCCCTGTCTCTAAACAGTCAAGCCCAAGC	3720
DB	3661		
QY	3721	CTTCCAGGGCCCTGTCTCTGCTCAGGTCCCTGAGCAGAGGAGATATGGCTCA	3780
DB	3721		
QY	3781	GGAAATGCCATGCTGCTTACATGTGTGTACTAGAGATATCCATAGTCTCTGGAGC	3840
DB	3781		
QY	3841	CTCTTAGGGTCTTTGGCTGGGTTGGGAGAACTTTACTCTCCCTCATATTCGTCATCA	3900
DB	3841		
QY	3901	CATACAGGGAGCTTGAGACACAGCTCTGTGTATTAAGGACACGTGTGAAGTCTGTGTGT	3960
DB	3901		
QY	3961	GT	4020
DB	3961		
QY	4021	GCCTAGTGTACCTCCGTGGCAGGATGGTGTAAACAGTGTATGAGTGCAGCTCTTTGAGCT	4080
DB	4021		

QY	4081	TTTAGCCTTGTCACTAGCCTTTTATTATACATCTCGAGAGTGTCTCAGTGTGTGTCTAC	4140
DB	4081		
QY	4141	AAAGACAGCGCCAGCCCTCTTCTGTCTGAGCTGTGTGAGCAGAGTGCCAGTCAACTCCAC	4200
DB	4141		
QY	4201	GGGCTTATGACACCGCAGCCTACCAAGCATGGCTGTCTATCCCTCTGGCTCTTAAGGTC	4260
DB	4201		
QY	4261	CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTTGTAGCATCTCTGTACCTAAT	4320
DB	4261		
QY	4321	TTGTAATCTGGAAAGTGCTGGTGGGAAATCTTCTTTCGCAACCTGTCTCTGCTC	4380
DB	4321		
QY	4381	CTTCTCTTCAATTTGTCTGGTGATCTGTCTCATGTCTCATGTCTCATGTCTCTGGGCC	4440
DB	4381		
QY	4441	CTTCTCTTCCCATGATGCCCTGATTTCTCACTGCTGTCTTCAATTTCTGTCTGCTCATG	4500
DB	4441		
QY	4501	CTTGTCTTTATGTCTGTGTCTCTGCTCCCTGAGTTCAACCTATGCAACCTTTCTTAAACA	4560
DB	4501		
QY	4561	ACATGACTACCTCATGTCTGTCTGAGACCATAGTGTGACCTGGGTGCTCCACAGTCTCC	4620
DB	4561		
QY	4621	CTGCCAACCGCTTCTTGGGAGATGAGCCCACTCAGTAGATCTCGAAGAGAGCCCTTG	4680
DB	4621		
QY	4681	TGGCTTGTCTGGCTGCCCTCTTGGTGTGTGAGATGAGAGGTTTCTATGGAAGAGAT	4740
DB	4681		
QY	4741	GAGTCAAGCTGCAAGGGGAAACCCCAAGAGGGGTAGGAGTGAACCAAGAGGCTGA	4800
DB	4741		
QY	4801	AAAAAATGGCTGCCACCCATCTGCACAGAGAGATGGGTGTGTGTGTGTGTGTGTGTGT	4860
DB	4801		
QY	4861	CTGGCTGAAACTGAAAGGGGTGAGGAGGGGAGCTACTGGGGCTGCTCAGTCTCC	4920
DB	4861		
QY	4921	CTGACCTTGAAGCCCTGAACTGGCTTCAAGATGAGCAAGAGTTTCTCTCAAGATGCTGT	4980
DB	4921		
QY	4981	AAAGGAAGTCTTTGATAGGAAAAGGGGGCTGGCTCATTTTATTTTATCTTTTCTTACA	5040
DB	4981		
QY	5041	CTGAATCCCAAAATCATCTTACCAAAAGGGGCAAGCCCTGATCTGTATTTCTGAGTCAAC	5100
DB	5041		
QY	5101	AAAGGCCATGCAATCTCTGTGTGTCTCACTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5160
DB	5101		
QY	5161	GTGGCATCTGTGCTCTTGTGTGCACATCTTCTATTTCAACTGGCTGGGCACATCAAGTGT	5220

Db 5161 GTGGCATCTGCTCTTCTGTCACATCTTCTTATTTCACTGGCTGGCACATCAAGTGT 5220
Qy 5221 AACTCTGGCTCTGGGCCAAGTTAGAAATAACAGCTATTTTCCCTTTATTTATTTT 5280
Db 5221 AACTCTGGCTCTGGGCCAAGTTAGAAATAACAGCTATTTTCCCTTTATTTATTT 5280
Qy 5281 TTTTATTTTATTTATGCTCTTTCAGTGGAGTTGTAGCTTCTGAAGCGCTGTGTATT 5340
Db 5281 TTTTATTTTATTTATGCTCTTTCAGTGGAGTTGTAGCTTCTGAAGCGCTGTGTATT 5340
Qy 5341 AGCCTGTGTGTCACCTCATGTTTGGACCCACCCACCAATTCCTTCTCCCTCTTCAGC 5400
Db 5341 AGCCTGTGTGTCACCTCATGTTTGGACCCACCCACCAATTCCTTCTCCCTCTTCAGC 5400
Qy 5401 CAGCCTATGATAAATAAGATTTATTAATGCTGGCTTGTATCTCATTTAAAGACAGAT 5460
Db 5401 CAGCCTATGATAAATAAGATTTATTAATGCTGGCTTGTATCTCATTTAAAGACAGAT 5460
Qy 5461 TGTCACTTGAACCTCTTCTATAGCAATTCAGAGTGGCCAGCGCCAAACACCCGATGTTT 5520
Db 5461 TGTCACTTGAACCTCTTCTATAGCAATTCAGAGTGGCCAGCGCCAAACACCCGATGTTT 5520
Qy 5521 CTTTCACTGCTGAAGGTCAGAGGCTCATTTTGTGTTTCTGGTTAGATCTTTTCTCTCC 5580
Db 5521 CTTTCACTGCTGAAGGTCAGAGGCTCATTTTGTGTTTCTGGTTAGATCTTTTCTCTCC 5580
Qy 5581 TTGCTTGAATCAAAATAACCGTTTAAACAGTAGGCTCTTAGCATCACACCATAGTAT 5640
Db 5581 TTGCTTGAATCAAAATAACCGTTTAAACAGTAGGCTCTTAGCATCACACCATAGTAT 5640
Qy 5641 TCTCATGCTTCTGTTTAAACAGCACTTGAAGTCTGGGTTTAAATTAAGTGCACAA 5700
Db 5641 TCTCATGCTTCTGTTTAAACAGCACTTGAAGTCTGGGTTTAAATTAAGTGCACAA 5700
Qy 5701 TGAGCAATTTATAACCAATAGGCTGGGTGAAATTTGTTCTCAAAAGCAATTAAGTAA 5760
Db 5701 TGAGCAATTTATAACCAATAGGCTGGGTGAAATTTGTTCTCAAAAGCAATTAAGTAA 5760
Qy 5761 TAAATCTGATCTGCTTAACTCAAGTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 5820
Db 5761 TAAATCTGATCTGCTTAACTCAAGTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 5820
Qy 5821 TTATATGATTTAGGCTTCTGAGTAACTGGGAGTGTAGCTTGTGACTTTGTAGCACC 5880
Db 5821 TTATATGATTTAGGCTTCTGAGTAACTGGGAGTGTAGCTTGTGACTTTGTAGCACC 5880
Qy 5881 AGGTCTTATAGGAAGTCTGTTGGCTTTTACAGGCAATAGTCCCTTTGTGCTTGGC 5940
Db 5881 AGGTCTTATAGGAAGTCTGTTGGCTTTTACAGGCAATAGTCCCTTTGTGCTTGGC 5940
Qy 5941 ATGGATGCTTAAAGTCTTGTGAGTCTCATTTAAGAAATTCCTTTCTCGAAGCATGACA 6000
Db 5941 ATGGATGCTTAAAGTCTTGTGAGTCTCATTTAAGAAATTCCTTTCTCGAAGCATGACA 6000
Qy 6001 GTGTATGCAATPACTTACATGCTCACTGTTTACCTGGCTAGTTTGTGCTGGGTATT 6060
Db 6001 GTGTATGCAATPACTTACATGCTCACTGTTTACCTGGCTAGTTTGTGCTGGGTATT 6060
Qy 6061 AATTGACATTTCCAGCATCATGCTTCTCTTACAAATATGATATTTTATTTGTTACAC 6120
Db 6061 AATTGACATTTCCAGCATCATGCTTCTCTTACAAATATGATATTTTATTTGTTACAC 6120
Qy 6121 TAAGTGTGTATCATGATCTGCTCCCTGTTAAAGAAATTAATAAATTAATTTTCAGAC 6176
Db 6121 TAAGTGTGTATCATGATCTGCTCCCTGTTAAAGAAATTAATAAATTAATTTTCAGAC 6176

RESULT 2
AAI67204
ID AAI67204 standard; DNA; 3753 BP.
XX
AC AAI67204;

XX 11-FEB-2002 (first entry)
XX Nucleotide sequence of GSK gene Id 27142.
XX Peptide hormone; antidiabetic; anorectic; antianorectic; antiaesthetic;
XX anidipressant; nootropic; neuroprotectant; hypotensive; hypertensive;
XX cytosatic; cerebroprotective; vasotropic; human; ds.
XX Homo sapiens.
XX WO200172961-A2.
XX 04-OCT-2001.
XX 22-MAR-2001; 2001WO-US09226.
XX 24-MAR-2000; 2000US-192158P.
XX 28-MAR-2000; 2000US-192668P.
XX 27-APR-2000; 2000US-200166P.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RP, Xiang Z, Kabnick KS;
XX Lai Y;
XX WPI; 2001-639223/73.
XX P-P8DB; AAG65914.
XX Isolated polypeptides, which may be peptide hormones, which are
XX identified by high throughput genome-based biology which identifies
XX genes and gene products as therapeutic targets for treatment of
XX diseases such as diabetes and cancer
XX Claim 2; Page 55-56; 99pp; English.
XX The invention provides polypeptides (AAG65986-65918) which may be peptide
XX hormones (including insulin, growth hormones, chemokines, cytokines,
XX neuropeptides, integrins, kallikreins, lamins, melanins, natruiretic
XX hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,
XX secretogranins, selectins, thromboglobulins, thymosins) identified by
XX high throughput genome-based biology and polynucleotides (AAI67176-67208)
XX encoding them. The polypeptides can be expressed by standard recombinant
XX methodology. The polypeptides are useful in the treatment of disease such
XX as diabetes, breast-, prostate-, colon cancer and other malignant tumors,
XX hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,
XX asthma, manic depression, dementia, delirium, mental retardation,
XX Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental
XX or sexual development disorders, and dysfunctions of the blood cascade
XX system including those leading to stroke. The polynucleotides may be used
XX as diagnostic reagents through detecting mutations in the associated gene
XX and for chromosome localization and for tissue expression studies. The
XX polypeptides and polynucleotides may also be used as vaccines.

SQ Sequence 3753 BP; 714 A; 1220 C; 1152 G; 667 T; 0 other;
Query Match 45.4%; Score 2804.2; DB 22; Length 3753;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 3171; Conservative 0; Mismatches 568; Indels 6; Gaps 2;
Qy 3 GGCGGGGCGGACACGGGCGGCTCTCTGCTGCTGACCTTCTGCTGCTGCTGCTGCGCG 62
Db 9 GGGGGAACCGCGCGCGCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 68
Qy 63 CGGGGAGCTGCTGCTTGGCCCCCAGGAGACCACTGTCAAGCTGCTGCTGCTGCTGCTGCTGCT 122
Db 69 CGGGGAGCTGCTGCTTGGCCCCCAGGAGACCACTGTGAGCTGCTGCTGCTGCTGCTGCTGCT 128
Qy 123 GCAAGTGTATCTGGGCGCTGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 182
Db 129 GCAAGTGTATCTGGGCGCTGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 188

Qy	183	TGCTGGGCTCCGACCAAGGCTGATGGAGCAAGGATGGAGACACTGTACTAGAGCATGA	242	Db	1266	GGTGTGGCGAGGGGCTGCCAGCGCCCCACGCGGGGTCACTGTCTAGCCACTGAGCAG	1325
Db	189	CGCTGGACCCCCACACAGGCTGACCTGGAGCAAGGATGGGACACACCTGTGGAGCACGA	248	Qy	1323	CTCCCTCTGTGTGTGGGCTGGGAGCGGCTGGAGTTGCACAGCGAGCAAAATCATTTGGCTT	1382
Qy	243	GAACCTGCACCTGCTACCAATAGCTCCCTGTGGCTGTCTCACTCCCTAGAGCAAGAA	302	Db	1326	CTCCCTGTGTGTGGGCTGGGAGCGGCTGGAGTTGCACAGCGAGCAGATCATCGGCTT	1385
Db	249	CCACTTACACCTGTCCCAATGGTTCCCTGTGGCTGTCCAGCCACTAGCACCAATGG	308	Qy	1383	CTCTCTTCACTACCAAAAGGCAAGGGGAGTGACAAATGTGGAGTACCAAGTTTGCAGTAAA	1442
Qy	303	CAGCGATGATGAGAAAGCTCTTAGGATCTGGAAGGTCACTTGAGGGCAGCTATTTCGTGTCT	362	Db	1386	CTCTCTCCACTACCAGAGGCAAGGGCATGGCAATGTGGAAATACCAAGTTTGCAGTGA	1445
Db	309	CAG---TGACAGTCACTCTCAGGCTGTGGGGGTCAATTGAAGGCAACTATTTCGTGTCT	365	Qy	1443	CAATGACACCAAGAGCTGCGAGTTTCGGGACCTGGAACTCCAAACCGAATATGAGTTCTA	1502
Qy	363	GGCCCAAGCCCGCTAGGAGTGTGGCCAGCCAGGTTGTGTGGTCAAGCTTGGCCACACT	422	Db	1446	CAACGACACCAAGAACTACAGTTTCGGGACCTGGNACTCCAAACCAAGATATGAGTTCTA	1505
Db	366	AGCCCAAGCCCGCTCGGAGTGTGGCCAGCCAGACTGTCTGTCAAGCTTGCACACT	425	Qy	1503	CGTGTGGCTTACTCCAGCTGGGGGCGAGCGAAGCTCCAGCCAGCGCTGGTGCATAC	1562
Qy	423	CGAAGACTTCTCTGTGACCCCGAGTCCAGATGTGGAGGAGAAAGGAGCAGCAGCTT	482	Db	1506	CGTGTGGCTTACTCCAGCTGGGAGCCAGCGCACTCCACCCAGCACTGGTGCACAC	1565
Db	426	CGCAGACTTCTCTGTGACCCCGAGTCTCAGACCGGTGGAGGAGAAAGGAGCAGCTCGCTT	485	Qy	1563	ACTGACAGATGTCCCGAGCGACACCCAGCTTACTTGTCCAGCCCAACCCCTCGGA	1622
Qy	483	TGAATGCCACCAAGGGCTCCAGCCCGCATCTTCCCAAGTGGCTCCTCCAGATCCTAGATGT	542	Db	1566	ACTGATGATGTCCCGAGTGCAGCACCCAGCTCTCCCTGTCCAGCCCAACCCCTCGGA	1625
Db	486	TGAGTGCACATTTGAAGGGCTGCAGCTCCCATCTATTCTTGGGAGAGAGCAGGTGAC	545	Qy	1623	CAATGAGGTGGATGGCTGCCCTCCCTCCAGCTGAGCAATGAGCAGGTGCTGAAAGTA	1682
Qy	543	CGTGCCTGAGGAGCCCGGCTCATCACTCTTCCCAAGTGGCTCCTCCAGATCCTAGATGT	602	Db	1626	CATCAGGGTGGCTGGCTGCCCTGCCCTCCAGCTGAGCAATGGCAGGTGGTGAAGTA	1685
Db	546	ATTGCCCTGAGGAGCCCTCGCTCATCGTCTTCCCAAGCGGCTCTTTCAGATCTCGATGT	605	Qy	1683	CAAGATAGAGTACGGTTTGGGGAAGGAGATCAGATTTTCTCTACTGAGGTGCGAGGAAA	1742
Qy	603	CAAGGACAGTGTGAGGCTCTTACCGCTGGCTGGCCACCAATTCAGCCCGCCCAAGATT	662	Db	1686	CAAGATAGAAATCGGTTTGGGAAAGGAGATCAGATTTTCTCTACTGAGGTGCGAGGAAA	1745
Db	606	TCAGGAGAGTGTATGAGGCGCCCTTACCGCTGGCTGGCCACCACTCAGCTCGCCAGCACTT	665	Qy	1743	TGAGACACAACTTACGTTAAACTCACTTCAGGCCAAACAAAGGTGACCGAGTCCGATTTTC	1802
Qy	663	CAGCAGGAGGCTCTCACTGTGGCTCAGAGGCTCTTGGAGGCTTACAGGGGGCA	722	Db	1746	TGAGACACAGCTTATGTCTGACTCGCTTCAGGCCAAACAAAGGTGATTCGAGTACGATTTTC	1805
Db	666	CAGCAGGAGGCTCTCACTGTGGCTCAGAGGCTCTTGGAGGCTTACAGGGGGCA	725	Qy	1803	AGCTGGCACTGGCTGTGCTATGAGTCCCTTCTCAGTGGATGCGACACAGCACTCGG	1862
Qy	723	GGATGTGGTCAATTTGGAGCCCGCAGAGAACACACAGGTAGTGTCTGAGCAGAAATGTAGT	782	Db	1806	GGCTGGTAAGCAGCAGCGGCTTCGGGCGCCCTCCAGTGGATGCAATGAGGTCAGGCGCCAG	1865
Db	726	GGAGTGGTCAATTTGGAGCCCGCAGAGAACACACAGTGGTGTCTGGCCAGATGTGTGT	785	Qy	1863	TGTGCACACACAGAGCCATGTTCCCTTTGCCCCCTGCAGNAATTCAGAGGTGAGGGCAAGAT	1922
Qy	783	GATGAGTGGTGGCTCTGCTGACCCCGCTTTTGTGTCTGGTTCGAGCAGAGTGG	842	Db	1866	TATGCACAAACAGAGCCATGTCCTTTTGGCCCCCTGCAGAGTTGAGAGGTGAGGGCAAGAT	1925
Db	786	GATGGAATGTGGGCTCAGCTGACCCCGCTTTTGTGTCTGGTTCGAGCAGAGCAG	845	Qy	1923	GGAGTCCCTGTGTGTATGTCAGCGCGCCCTCACTCCAGCCAGATCTCTGAGTACAA	1982
Qy	843	AAAGCTTATCCACCGATGCTATGTTCTGGGCGGACCAATCTACTCATCCAGAGCC	902	Db	1926	GGAGTCCCTGTGTGTATGTCAGCGCCCGCTCACTCCAGCCAGATCTCTGAGTACAA	1985
Db	846	GAAGCCCAATCCACAGATGCTATGCTCTGGGCGGACCACTACTAATGCCAACGC	905	Qy	1983	ACTCTACTGGGAGAGGTGGGAAACAGAGGAGGAGGAGTGTGACCCGCCCCCGAGGGG	2042
Qy	903	GCAGCTTCGGCACTCTGGAGTCTATGTCTGCGGAGCCCAAGCCCTCAACGCGTGA	962	Db	1986	ACTATATTGGCGGAGGTGGGGCTGAGGAGGAGGCAATGGCGATCGCTGCCAGGGG	2045
Db	906	GCAGCTTCGGCACTCTGGAGTCTATGTCTGCGGCGGACCAAGCCCGCGAGCGG	965	Qy	2043	TGCTGGAGATCAAGCTTGGGACGTGGGCGCGTGGGCTGAGAGAGAGAGTGAAGCAGTA	2102
Qy	963	CGCCACTTCGGGCTGAGCTCGAGTGTGTCGCCCGAGCCATCTGCGAGGACCCGA	1022	Db	2046	CGTGGAGACCAAGGCTTGGGATGTGGGGCTTGTCCGGCTCAAGAGAGAAAGTGAAGCAGTA	2105
Db	966	CGCCACTTCGGGCTGAGCTCGGCTGCTGGGCGGCTCCCGCCATCACTCAGGCGCCGA	1025	Qy	2103	TGAACTGACCCAGTTAGTCCCTGCGCAGGCGCTGACGAGGTGAAGCTCTGAGCTTTCAACAA	2162
Qy	1023	GGCGCTTCGGGAGCGGGGCGAGCAGCGCGCTGTGTGTGGGCGGCTCGGGGAGCC	1082	Db	2106	TGAGCTGACCCAGCTAGTCCCTTGGCGGCTGTGACGAGGTGAAGCTCTGAGCTTTCAACAA	2165
Db	1026	GGCGCTTCGGGAGCGGGGCGAGCAGCGCGCTTGTGTGGGCGGCTCGGGGAGCC	1085	Qy	2163	ACACGAGGAGCGGCTACGCTGTGTGTGGAGGGGCAAGCGAGAGAGGCGCCCAAGCAG	2222
Qy	1083	ACGGCCCGCTGCTGCTGAGCAGCGGAGTCCGTTGCGGACCCCAATGGGGCGCTCAA	1142	Db	2166	ACATGAGGATGGCTATGACAGCAGTGTGGAAAGGGCAAGACGAGAGGGCGCCGACCA	2225
Db	1086	CGGGCCAGCGCTGGCTGCTGAGCAGCGGGGCGCTGCGGCGCCCAAGGGGCGCTCAA	1145	Qy	2223	CCTGCTATTCAGAGGGGCGACCGCTGCTCTGCGCCATGTGCGAGAGTCAACAG	2282
Qy	1143	GGTCCAGGGCGGTGGGCGAGCTTGTGATCACTCAATCGGCTGCGAGGAGCTGGCTA	1202	Db	2226	CATGCTATTCAGAGGGGACCACTCCCTGCTCCAGCCCACTGCTCAATGCGGAAATCAACAG	2285
Db	1146	GGTCCAGGGCGGTGGGCGAGCTGCTGATCAACACAGATCGGCTTGCAGGAGCGCGCTA	1205	Qy	2283	CTCCACTTCCATTTGGCTTGGTGAAGAGGAGGAGCTTTTACACTGTCTCAAGATTGTCAA	2342
Qy	1203	CTACCAAGTGGTGGTGAAGAAACAGCGGGGAAATGCGCTGTGCGGCTCGGCGCTGGCGGT	1262	Db	2286	CTCCACTTCCATTTGGCTTGGTGAAGAGGAGGAGTTCACACAGTCAAGATTGTCAA	2345
Db	1206	CTACCAAGTGGTGGTGAAGAAACAGCGGGGAAATGCGCTGTGCGGCTCGGCGCTGGCGGT	1265	Qy	2343	CTACACTGTGACGCTTGGCGGCTCGGAGTGTCTTCTGCTGCTCACCTACTATAC	2402
Qy	1263	AGTGGTGGGAGGGGCTGCCAGCGCCGACTCGGGTCAACAGCCAGCGCGCTGAGCAG	1322				

Db 2346 CTACACTGTGCGCTTCAGCGCCCTCGGGGGCTCAGGAATGCTCCCTCGGTGACCTATTACAC 2405
 QY 2403 CAGCTCTGAGAGAGACATTTCTCATTTGGCGGCTGAAACCATTTTACCAAGTACGAGTTTGC 2462
 Db 2406 CAGTTCTGAGAGACATCTCTCATTTGGCGGCTTGAAGCAATTCACCAATACGATTTGC 2465
 QY 2463 GGTACAGTCCACGAGTGGATATGATAGGGCCCTTTTGGCTCGGTGGTAGAACCTCCAC 2522
 Db 2466 AGTGCAGTCTACCGCGGTGGACATGATGGGCTTTTGGCTCTGTGGTGGAGCGCTCCAC 2525
 QY 2523 CTGTGCTGACCGGCTTCAACACCTCTCTGACCTGCGCTGAGCCCTGACACCATC 2582
 Db 2526 CTGTGCTGACCGGCTTCAACACCTCTCTGACCTGCGCTGAGCCCTGACACCATC 2585
 QY 2583 CACCGTTCCGTTACACTGTGTGTCCTCCACGAGGCCCAATGCTGAGATTTGAGATCT 2642
 Db 2586 CACGTTCCGTTGCTGCTGCTGTCCTCCACGAGGCCCAATGCTGAGATTTGAGATCT 2645
 QY 2643 AATTCTACAGAAACACACACACCGAGCCGAAACACAGTGGACACTGCTCACACAGA 2702
 Db 2646 GATCTGTACAGCAGCAACACACACGAGCTGAGCACCAAGTGGACCTGCTCACACGCA 2705
 QY 2703 GGGAAACATCTTCAGTGCAGAGGTCCATGGCTTAGAGTGAACACTCGGTATTTCTTCAA 2762
 Db 2706 GGGAAACATCTTCAGTGCAGAGGTCCATGGCTTAGAGTGAACACTCGGTATTTCTTCAA 2765
 QY 2763 GATGGAGCGCCGACACAGAGTGGGCTGCGGCTTTTCCGCTTTCAGAGTGTGATTAC 2822
 Db 2766 GATGGAGCGCCGACACAGAGTGGGCTGCGGCTTTTCCGCTTTCAGAGTGTGATTAC 2825
 QY 2823 TCTGAAGAGACATTTCTAGACTCTCTGAGTGCACGCGCTCAGCGGCATCATCGTGG 2882
 Db 2826 GCTCCAGGAGAGCTGTGAGACTCGCTGGACATGACTCAGTACGCGGCATCATCGTGG 2885
 QY 2883 TGTCTGCTGGGCTTCTGCTCTGCGCTGCGATGTGTGCTGCGCTACGACAAAGCTC 2942
 Db 2886 TGTCTGCTGGGCTTCTGCTCTGCGCTGCGATGTGTGCTGCGCTGCGCGCAGGCC 2945
 QY 2943 CCACAGGGAAGCCCTTCCGAGTGTCTCTCAGGACCCGACGACCCGAGGAACCCAGCGCTCTA 3002
 Db 2946 CCACAGGGAATCCCTCCAGGCTGTCTCTCAGGACCCGACGACCCGAGGAATCCGCGCTCTA 3005
 QY 3003 CACAGAGCTCGGCTTGGGCTCCAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3062
 Db 3006 CTCAGAGCTCGGCTTGGGCTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3065
 QY 3063 GCATCTGCTCCGAGATTTGTCTCCACACCTCAGATGTGGAAGCAAGGCTGAAGT 3122
 Db 3066 GCACCCCATCCGAGGACTTGTCTCCGCTCAGACCTCAGACGTTGAGGACAGGGCTGAAGT 3125
 QY 3123 ACACAGCTTATGGGTGCGAGTGTCTCAGATTTGCGGGGCTCCTCAGAGGAAGATCTC 3182
 Db 3126 GCACAGCTTATGGGTGCGAGTGTCTCAGATTTGCGGGGCTCCTCAGAGGAAGATCTC 3185
 QY 3183 CTGGGCTCAGGAGGGGACCAAACTGGGAGGCTCTGGGAGGCTGTGAGCTGCCCA 3242
 Db 3186 CTGGGCTCAACCAAGCGGCTGAGCTGGGCTGGTCTCTGGGAGGCTGTGAGCTGCCCA 3245
 QY 3243 GGGTATGTCTCAAGCGGCTCTGACCGGCTCTGCTGCTTCCAGCGGGAACCGGCA 3302
 Db 3246 GG---CAGGCCCCCGGGCTCTGACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3302
 QY 3303 GACACTGTCTGCAAGCGCTTGTATGACGGCATTAAGAGCAACGAGGAGGAAGGCC 3362
 Db 3303 GACGCTGTGCTGCAAGGCTTGTATGACGGCATTAAGAGCAACGAGGAGGAAGGCC 3362
 QY 3363 GTCCCGAGCTGCAGGAATCAGGTGGAAGCTCAGGTCTATTGTCCACTCCGACTCGGTGTC 3422
 Db 3363 ACCCCAGGCTGCAGGAACAGGTGGAAGCTCAGGTCTATTGTCCACTCGGTGTC 3422
 QY 3423 ATCCAAAGAGTCTCTGACCTCCACCTCCAGGCTGAGGAGGAGGAACCACTGACTGC 3482
 Db 3423 ATCTAAGGAGGACCTGACCTCCATCTCCAAAGCTTGGAGGCTGAGGAGGAGGAGGAGGAGG 3482

QY 3483 AGAGACTGTGCTTCCAGCTCTGGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTGGG 3542
 Db 3483 AGAGCTCTGATCTCATCTCGGTGTTGGGATCCAGGGCAGGGGAGGCTGGCTGGA 3542
 QY 3543 CAGGAGCTGGAGGGTGCACCAACCAACCAAGTGGGCGAGAGGCTCAGCTGCTTGGC 3602
 Db 3543 CAGGAGTGGGAGGCTGTGAGCTGGCAGCCCGGCGCAGACAGACTTACCTGCTTGC 3602
 QY 3603 AGAAGCAGCAGTCTCTGCTCTGCTCAGACTCCAGCCAGCAGCTGCTATAGAGGA 3662
 Db 3603 AGAGCAGCAGTCTCTGCTCTGCTCAGACTCCAGCCAGCAGCTGCTATAGAGGA 3662
 QY 3663 GGGCCCTGGGAAAAGCTGCCAGCCCAAGCCCTGTGTCTCTAAGCAGTCAAGCCAAAGCT 3722
 Db 3663 GACCCCTGGAGATAGCTGCCAGCTCAATCCCTGCTCTAGGAGCCAGCCAGGCT 3722
 QY 3723 TCCAGGCGCTCTGCTCTGCTCTGCT 3747
 Db 3723 GCCAGATCCCGCTCTCTCTCTCT 3747

RESULT 3

AADI0022

ID AADI0022 standard; cDNA; 2796 BP.

XX AC AADI0022;

XX DT 12-SBP-2001 (first entry)

XX Mouse Nope (neighbour of punc ell) extracellular domain cDNA.

XX Mouse; Nope; neighbour of punc ell; cytostatic; neuroprotective; vaccine;
 gene therapy; cerebroprotective; colonic cancer; mental retardation;
 tumour suppressor; chromosome 9; transgenic animal; genetic disorder;
 obesity; Bardet-Biedl syndrome; autosomal recessive disorder; retinitis;
 polydactyly; pigmentosa; hypogonadism; hypertension; diabetes mellitus;
 renal anomaly; cardiovascular anomaly; extracellular domain; es.
 XX Mus musculus.

FH Key Location/Qualifiers

CDS 1..2796

FT /*tag= a

FT /product= "Mouse Nope extracellular domain"

FT /note= "CDS does not include start and stop codon"

FT /partial

XX WO200149714-A2.

XX 12-JUL-2001.

XX 26-OCT-2000; 2000WO-US29699.

XX 04-JAN-2000; 2000US-0174496.

XX 19-MAY-2000; 2000US-0205789.

XX (NEUR-) NEUROSCIENCES RES FOUND INC.

XX Salbaum JW;

XX WPI; 2001-441846/47.

XX P-FSDB; AAB05252.

XX Murine Nope polypeptides and nucleic acids useful for preventing,
 diagnosing and treating colonic cancer and Bardet-Biedl syndrome -
 Claim 11; Page 77-81; 99pp; English.

XX The present invention relates to Nope (neighbour of punc ell) which is

CC used in the prevention, treatment and diagnosis of diseases associated
 CC with inappropriate Nope expression such as cancers especially colonic
 CC cancer and genetic disorders, as Nope is thought to be a tumour

CC suppressor. Nope gene is located on chromosome 9 and is used in gene
CC therapy. Nope is used as vaccine. Nope gene may be administered to treat
CC diseases by rectifying mutations or deletions in a patient's genome that
CC affect the activity of Nope by expressing inactive proteins or to
CC supplement the patients own production of Nope polypeptides. Nope gene
CC is used to study the expression and function of Nope polypeptides and
CC their role in metabolism through the creation of transgenic animal
CC models. The anti-Nope antibodies and Nope antagonists may also be used
CC to down regulate Nope expression and activity for the treatment of
CC Bardet-Biedl syndrome which is an autosomal recessive disorder
CC characterised by mental retardation, obesity, polydactyly, retinitis
CC pigmentosa and hypogonadism. Patients with Bardet-Biedl syndrome have a
CC high incidence of hypertension, diabetes mellitus and renal and
CC cardiovascular anomalies. The present sequence is mouse Nope (neighbour
CC of punc ell) extracellular domain cDNA.

XX Sequence 2796 BP; 591 A; 849 C; 821 G; 535 T; 0 other;

Query Match 45.3%; Score 2796; DB 22; Length 2796;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2796; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 64 GGGAGCTGCCATTGGCCCGGAGGACAACTGTCAAGCTGAGCTGTGTATGAGGGACCCCTG 123
Db 1 GGGAGCTGCCATTGGCCCGGAGGACAACTGTCAAGCTGAGCTGTGTATGAGGGACCCCTG 60
Qy 124 CAATGATCTTGGCCCTGAGCAGCTGTGTGCTGCACTGTGCGCTTTGGGGCTACAGCT 183
Db 61 CAATGATCTTGGCCCTGAGCAGCTGTGTGCTGCACTGTGCGCTTTGGGGCTACAGCT 120
Qy 184 GCTGGGCTCCGACAGGCTGACATGAGCAGCAAGATGAGACACTGTACTAGAGCATGAG 243
Db 121 GCTGGGCTCCGACAGGCTGACATGAGCAGCAAGATGAGACACTGTACTAGAGCATGAG 180
Qy 244 AACCTGCACCTGTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAAGAC 303
Db 181 AACCTGCACCTGTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAAGAC 240
Qy 304 AGCGATGATGAGGAAGCTCTTAGGATCTGGAAGTCACTGAGGGCAGCTATCTGTCTG 363
Db 241 AGCGATGATGAGGAAGCTCTTAGGATCTGGAAGTCACTGAGGGCAGCTATCTGTCTG 300
Qy 364 GCCACAGCCGCTAGAGGTGTGGCCAGCAGGTGTGTGGTCAAGCTTGGCCACACTC 423
Db 301 GCCACAGCCGCTAGAGGTGTGGCCAGCAGGTGTGTGGTCAAGCTTGGCCACACTC 360
Qy 424 GAAGACTTCTCTGCAACCCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGCTTT 483
Db 361 GAAGACTTCTCTGCAACCCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGCTTT 420
Qy 484 GAATGCCACAAAGGGCTTTCCAGCCCGCATATTACTTGGGAAAGGACAGGTGACC 543
Db 421 GAATGCCACAAAGGGCTTTCCAGCCCGCATATTACTTGGGAAAGGACAGGTGACC 480
Qy 544 GTGCTGAGGAGCCCGCTCATCTCTCCCAAGTGGCTCTCCAGATCTAGATGTC 603
Db 481 GTGCTGAGGAGCCCGCTCATCTCTCCCAAGTGGCTCTCCAGATCTAGATGTC 540
Qy 604 CAGGACAGTGTAGGCTCTTACCGCTGTGGCCACCAATTGAGCCCGCCCAAGATTTC 663
Db 541 CAGGACAGTGTAGGCTCTTACCGCTGTGGCCACCAATTGAGCCCGCCCAAGATTTC 600
Qy 664 AGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGCTTTTGGAGGCTTACAGGGGGCAG 723
Db 601 AGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGCTTTTGGAGGCTTACAGGGGGCAG 660
Qy 724 GATGTGTCATGTGGCCCGCAGGAAACCAAGTGTGTGGACAAATGTAGTG 783
Db 661 GATGTGTCATGTGGCCCGCAGGAAACCAAGTGTGTGGACAAATGTAGTG 720
Qy 784 ATGAGTGCCTGGCTCTGTGAGCCCAACCTTTTGTGTCTGGTCCGACAGATGGA 843
Db 721 ATGAGTGCCTGGCTCTGTGAGCCCAACCTTTTGTGTCTGGTCCGACAGATGGA 780

Qy 844 AAGCCTATCTCCACGAGTGTCTGTCGGCCGAGCAATCTACTCATCGCAGCGG 903
Db 781 AAGCCTATCTCCACGAGTGTCTGTCGGCCGAGCAATCTACTCATCGCAGCGG 840
Qy 904 CAGCCTCGGCACCTCTGGAGTCTATGTCTGCCGAGCCAAACAGCCCTCACGGTGTCTTC 963
Db 841 CAGCCTCGGCACCTCTGGAGTCTATGTCTGCCGAGCCAAACAGCCCTCACGGTGTCTTC 900
Qy 964 GCACATCGCGCTGTCTGAGCTCCAGTCTTGTCTGCCCGAGCCATCTCGCAGGACCCGAG 1023
Db 901 GCACATCGCGCTGTCTGAGCTCCAGTCTTGTCTGCCCGAGCCATCTCGCAGGACCCGAG 960
Qy 1024 GCGCTCTCGCGAGCGCGGCGCAGCAGCGCGCTCTGTGTGCGCGGCTCTCGCGGAGCCA 1083
Db 961 GCGCTCTCGCGAGCGCGGCGCAGCAGCGCGCTCTGTGTGCGCGGCTCTCGCGGAGCCA 1020
Qy 1084 CCGCCCGCTGTGCACTGTGCTGCAAGCGGATTCCTGTTCGAGCCCAATGGCGCGCTCAAG 1143
Db 1021 CCGCCCGCTGTGCACTGTGCTGCAAGCGGATTCCTGTTCGAGCCCAATGGCGCGCTCAAG 1080
Qy 1144 GTGAGGGCGGTGTGGCGGAGCTTGTGTCTATCATCTCAGATCGGCTGCGAGCAGCTGCTAC 1203
Db 1081 GTGAGGGCGGTGTGGCGGAGCTTGTGTCTATCATCTCAGATCGGCTGCGAGCAGCTGCTAC 1140
Qy 1204 TACCACTGTGCTAGCAGAAAACAGCGCGGAACTGTCTGTGCGGCTCTGCGCGGTA 1263
Db 1141 TACCACTGTGCTAGCAGAAAACAGCGCGGAACTGTCTGTGCGGCTCTGCGCGGTA 1200
Qy 1264 GTGCTGCGGAGGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCT 1323
Db 1201 GTGCTGCGGAGGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCT 1260
Qy 1324 TCCTCTGTGCTGTGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCT 1383
Db 1261 TCCTCTGTGCTGTGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCT 1320
Qy 1384 TCCTCTTCACTTACCAAAAGGCAAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1443
Db 1321 TCCTCTTCACTTACCAAAAGGCAAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1380
Qy 1444 AATGACACCAAGAGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCT 1503
Db 1381 AATGACACCAAGAGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCT 1440
Qy 1504 GTGCTGCGCTTACCTCCAGCTGTGGGGCGCAGCGAACTCTCAGCCAGCCCTGTGCTAC 1563
Db 1441 GTGCTGCGCTTACCTCCAGCTGTGGGGCGCAGCGAACTCTCAGCCAGCCCTGTGCTAC 1500
Qy 1564 CTGAGCAGTGTCCCGCAGCGCAGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1623
Db 1501 CTGAGCAGTGTCCCGCAGCGCAGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
Qy 1624 ATCAGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1683
Db 1561 ATCAGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
Qy 1684 AAGATAGATGTACGGTTTGGGGAGGAGATCAGGTTTCTCCACCGAGGTGCTGGAAT 1743
Db 1621 AAGATAGATGTACGGTTTGGGGAGGAGATCAGGTTTCTCCACCGAGGTGCTGGAAT 1680
Qy 1744 GAGACAACTTACGTTTAACTCACTTTCAGCCAAACAAAGTGTGCTGCTGCTGCTGCTGCT 1803
Db 1681 GAGACAACTTACGTTTAACTCACTTTCAGCCAAACAAAGTGTGCTGCTGCTGCTGCTGCT 1740
Qy 1804 GCTGGCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1863
Db 1741 GCTGGCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Qy 1864 GTGCACAAACAGAGCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1923
Db 1801 GTGCACAAACAGAGCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860

```
QY 1924 GAGTCCCTGGTGGTGTCTATGGCAGCGCGCCCTCTACCCACCAGATCTCTGGATACAA 1983
DB |||||
QY 1861 GAGTCCCTGGTGGTGTCTATGGCAGCGCGCCCTCTACCCACCAGATCTCTGGATACAA 1920
DB |||||
QY 1984 CTCTACTGGGAGAGGTGGGAACAGAGGAGGAGGAGTGGTACCGCCCCCAGGGGT 2043
DB |||||
QY 1921 CTCTACTGGGAGAGGTGGGAACAGAGGAGGAGGAGTGGTACCGCCCCCAGGGGT 1980
DB |||||
QY 2044 CGTGAGATCAAGCTTGGGACGTCGGGCGCCGTCGGCTGAAGAAGAAAGTGAAGCAGTAT 2103
DB |||||
QY 1981 CGTGAGATCAAGCTTGGGACGTCGGGCGCCGTCGGCTGAAGAAGAAAGTGAAGCAGTAT 2040
DB |||||
QY 2104 GAACTGACCCAGTGTAGTCTCTGGCAGCGCGTACGAGGTGAAGCTGTAGCTTTCAACAA 2163
DB |||||
QY 2041 GAACTGACCCAGTGTAGTCTCTGGCAGCGCGTACGAGGTGAAGCTGTAGCTTTCAACAA 2100
DB |||||
QY 2164 CACGAGGAGCGCTACGCTGTGTGTGAAGGCAAGACGAGGAGGCGCCACGCGCAGC 2223
DB |||||
QY 2101 CACGAGGAGCGCTACGCTGTGTGTGAAGGCAAGACGAGGAGGCGCCACGCGCAGC 2160
DB |||||
QY 2224 CTGCCTATCCAGAGGGGGCCACCGCTGCTCTCTGCCCCATGTCTCACGCGAGGTCAACACAGC 2283
DB |||||
QY 2161 CTGCCTATCCAGAGGGGGCCACCGCTGCTCTCTGCCCCATGTCTCACGCGAGGTCAACACAGC 2220
DB |||||
QY 2284 TCACCTTCATTTGGCTTGGTGGAGAGGAGCAGACTTTTACCTGTCAAGATTGTCTCAAC 2343
DB |||||
QY 2221 TCACCTTCATTTGGCTTGGTGGAGAGGAGCAGACTTTTACCTGTCAAGATTGTCTCAAC 2280
DB |||||
QY 2344 TACACTGTACGCTTCGGGCGCCCTGGGGGCTCAGGAATGCTTCCCTGGTCACTTATACC 2403
DB |||||
QY 2281 TACACTGTACGCTTCGGGCGCCCTGGGGGCTCAGGAATGCTTCCCTGGTCACTTATACC 2340
DB |||||
QY 2404 AGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTTACCAAGTACGAGTTTGG 2463
DB |||||
QY 2341 AGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTTACCAAGTACGAGTTTGG 2400
DB |||||
QY 2464 GTACAGTCCACGAGTGTGATATGAGTGGGCGCTTTGGCTCGCTGTGAGAACGCTCCACC 2523
DB |||||
QY 2401 GTACAGTCCACGAGTGTGATATGAGTGGGCGCTTTGGCTCGCTGTGAGAACGCTCCACC 2460
DB |||||
QY 2524 CTGCCTGACCGGCTTCAACACCTCTCTGTGACCTGCGCTGAGCGCCCTGACACCACTCC 2583
DB |||||
QY 2461 CTGCCTGACCGGCTTCAACACCTCTCTGTGACCTGCGCTGAGCGCCCTGACACCACTCC 2520
DB |||||
QY 2584 ACCGTTGGTTTACACTGTGTGTCTCCCGCCACGAGGCCCAATGTGTGAGATTGTGAGTATCTA 2643
DB |||||
QY 2521 ACCGTTGGTTTACACTGTGTGTCTCCCGCCACGAGGCCCAATGTGTGAGATTGTGAGTATCTA 2580
DB |||||
QY 2644 ATTCTTACAGCAACACCAACCCAGCGCCGAAACCAAGTGGACACTGCTCACCAAGAG 2703
DB |||||
QY 2581 ATTCTTACAGCAACACCAACCCAGCGCCGAAACCAAGTGGACACTGCTCACCAAGAG 2640
DB |||||
QY 2704 GGAACATCTTCACTGTCAGAGTGTCCATGCGCTAGAGAGTGACACTCGGTATTTCTCAAG 2763
DB |||||
QY 2641 GGAACATCTTCACTGTCAGAGTGTCCATGCGCTAGAGAGTGACACTCGGTATTTCTCAAG 2700
DB |||||
QY 2764 ATGGAGGCGCCACAGAGTGGGCTGGGCGCTTTTCCCGCTTCAGAGTGTGATTACT 2823
DB |||||
QY 2701 ATGGAGGCGCCACAGAGTGGGCTGGGCGCTTTTCCCGCTTCAGAGTGTGATTACT 2760
DB |||||
QY 2824 CTGCAAGAGACATTTCTCAGACTCTTTGGATGTGCAC 2859
DB |||||
QY 2761 CTGCAAGAGACATTTCTCAGACTCTTTGGATGTGCAC 2796
DB |||||
```

RESULT 4

ABK11101

ID ABK11101 standard; DNA; 3741 BP.

XX ABK11101;

AC ABK11101;

XX 05-JUN-2002 (first entry)

DT

XX

DE DNA encoding human NOV1 protein, homologue of NOPE/PUNC Ig proteins.

XX Human; NOVX-associated disorder; developmental disorder; blood disorder; endocrine disorder; vascular disease; gastrointestinal disorder; cancer; respiratory disorder; inflammatory disorder; reproductive disorder; neurodegenerative disorder; autoimmune disorder; infectious disease; cardiovascular disorder; cell signal processing; PUNC immunoglobulin; metabolic pathway modulation; NOPE; Ig; NOV1; gene; de.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..3741

XX /*tag= a

XX /partial

XX /product= "NOV1"

XX /note= "This sequence lacks a stop codon"

XX WO200206329-A2.

XX 24-JAN-2002.

XX 18-JUL-2001; 2001WO-US22709.

XX 18-JUL-2000; 2000US-218870P.

XX 18-JUL-2000; 2000US-218875P.

XX 18-JUL-2000; 2000US-218901P.

XX 24-JUL-2000; 2000US-220273P.

XX 26-JUL-2000; 2000US-220912P.

XX 27-JUL-2000; 2000US-221233P.

XX 28-JUL-2000; 2000US-221650P.

XX (CURA-) CURAGEN CORP.

PA Raetelli L, Shinkets RA, Zerhusen B, Malyankar UM, Padigar M;

PI WPI; 2002-179781/23.

DR P-PSDB; AAU77405.

XX Novel cytoplasmic, nuclear, membrane bound and secreted NOVX polypeptides, useful for treating developmental disorders, endocrine disorders, vascular disorders, infectious diseases and neurodegenerative disorders -

XX Claim 9; Page 8-9; 178pp; English.

XX The present invention relates to the isolation of novel human polypeptides referred to as NOV1, NOV2, NOV3, NOV4a, NOV4b, NOV5A, NOV5b, NOV6 AND NOV7, and the polynucleotide sequences encoding them. The NOVX polypeptides are related to NOPE, cadherin, interferon alpha-13, ADAM, ankyrin repeat-containing, transpanin or semaphorin polypeptides. The sequences of the invention are useful for identifying an agent (a cellular receptor or downstream effector) that binds to the NOVX polypeptide, or an agent that modulates its expression or activity. They are useful for treating or preventing NOVX-associated disorders such as developmental disorders, endocrine disorders, vascular diseases, gastrointestinal disorders, respiratory disorders, inflammatory disorders, blood disorders, reproductive disorders, neurodegenerative disorders, autoimmune and immune disorders, infectious diseases, cardiovascular disorders, cancers, and other disorders related to cell signal processing and metabolic pathway modulation. The present sequence encodes the human NOV1 protein.

XX Sequence 3741 BP; 708 A; 1209 C; 1150 G; 673 T; 1 other;

Query Match 43.4%; Score 2681.8; DB 24; Length 3741;
Best Local Similarity 83.7%; Pred. NO. 0;
Matches 3141; Conservative 0; Mismatches 583; Indels 30; Gaps 8;

QY 3 GGCGGGGGGACACGGCGCGGCTCTGGTGTGACCTTCTGCTGTCTGCGCGG 62

DB 9 GGCGGACCGCGCGCGGCGCGGCTCTCTGCGGTGACCTTCTGCTGTCTGCGCGG 68


```

2223 AGACATGCTATCCAGAGGGGACCAACCCCTGCTCCAGCCCAAGCTCCATCGGGAATCAA 2282
2280 CAGCTCCACTTCCATTGCTTGGTGGGAAGCCAGACTTTTACCACTGTCAAGATTGT 2339
2283 CAGCTCCACTTCCATTGCTTGGTGGGAAGCCAGACTTTTACCACTGTCAAGATTGT 2342
2340 CAACTACACTGTACCGCTTGGCCCTTGGGGCTCAGGAATGCTTCCCTGGTCACTACTA 2399
2343 CAACTACACTGTACCGCTTGGCCCTTGGGGCTCAGGAATGCTTCCCTGGTCACTACTA 2402
2400 TACCAGCTCTGAGGAAGACATTTCTATTGGCGGCTGAAACCAATTTTACCAAGTACAGATT 2459
2403 ---CAGTTCTGAGGAAGACATTTCTATTGGCGGCTTGAAGCCATTTTACCAATATCAGATT 2459
2460 TCGGTACAGTCCCAAGGAGTGGATATGATGAGGCTTGTGGCTCGTGGTGAAGACGCTC 2519
2460 TCGAGTGCAGTCTCAGCGGCTGGAATGATGGGCTTTCTGGCTCTGTGGTGGAGCGCTC 2519
2520 CACCTGCTGAGCGGCTTCAACACCTCTTCTGACCTGCGCTGAGCGCCCTGACACC 2579
2520 CACCTGCTGAGCGGCTTCAACACCTCTTCTGACCTGCGCTGAGCGCCCTGACACC 2579
2580 ATCCACCGTTCCGTTTACACTGTGTCCCGCCACGAGCCCAATGTGTGAGATTGTGAGTA 2639
2580 GTCCACGGTTCCGTTTACACTGTGTCCCGCCACGAGCCCAATGTGTGAGATTGTGAGTA 2639
2640 TCTAATTTCTACAGCAACCAACACACAGCCCGAACAACCAAGTGGACCTGCTCACCAC 2699
2640 TCTGATCTCTACAGCAGCAACCAACACAGCCCGAACAACCAAGTGGACCTGCTCACCAC 2699
2700 -----AGAGGGAACATTTTCAGTGCAGAGTCCATGCGCTTAGAGAGTGAACCTCGGTA 2753
2700 GCAGGGTGAAGGGAACATTTTCAGTGCAGAGTCCATGCGCTTAGAGAGTGAACCTCGGTA 2759
2754 TTTCTTCAAGATGGGAGCCCGCACAGAGTGGGCTTGGCCCTTTTCCCGCTTCAGGA 2813
2760 CTTCTTCAAGATGGGAGCCCGCACAGAGTGGGCTTGGCCCTTTTCCCGCTTCAGGA 2819
2814 TGTGATTTACTTCGAAGAGACATTTCTCAGACTTCCTTGGATGTGCACGCGCTCACCGGCAT 2873
2820 TGTGATCAGCTCCAGGAGAAGCTGTACAGCTCGTGGACATGCACTCAGTACACGGCAT 2879
2874 CATGTGGGTGTGCTGGGCTTCTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2933
2880 CATGTGGGTGTGCTGGGCTTCTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2939
2934 ACAAGCTCCACAGGAGCCCTTCCCGATTTGCTCTCCTCAGGACCCCAAGGAACCC 2993
2940 CCGAGCCCTCCACAGGGAATCTCTCCAGGCTGTCTCTCCACCGCCACCCCGGAATCC 2999
2994 AGCGCTCTACACAGAGCTCGGCTTGGGCTTCCAGTGTCTCCTGCTGCTGCTGCTGCTGCTGCT 3053
3000 CGCGCTGTACTCCAGAGCTCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 3059
3054 GTCCCTCGTGCATCTCTGCTCCAGGATTTGGTCCCAACCAACCTCAGATGTGGAAGACAA 3113
3060 GTCCCTTGTGACCCCTTCCAGGATTTGGTCCCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 3119
3114 GGCTGAGTACACAGCTTATGGTGGGAGTGTTCAGATTGCTGCGGGGCACTTCCAGAG 3173
3120 GGCTGAGTACACAGCTTATGGTGGGAGTGTTCAGATTGCTGCGGGGCACTTCCAGAG 3179
3174 AAAAGATCTCTGGGCTTCAGGAGGGGACCAAACTGGGAGGCTCTCTGGGAGGCTGTGA 3233
3180 AAAAGT---AAGTGTCTAACAAAGCGGCTGAGCTGGGCTGGTCTCTGGGAGGCTGTGA 3236
3234 GCTGCCCAAGGATGTGTTCAAGGCGGCTCTGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3293
3237 GCTGCCCAAG---CAGGCTTCCCGGCTGCTGACCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3293
3294 AACCGGCGAGACATCTGCTGCAAGCCCTGCTGATGATGAGCGGCAATAAGAGCAACGCGAG 3353
3294 AACTGGGCGAGCGCTTGTGCTGAGGTTCTCTGCTCTGATC-----AGGCAATGGGAG 3347
```

```

3354 AAAGAGCGCTTCCAGAGCTGCGAGGAATCAGGTGGAAGCTCAGGTCAATTGTCCACTCCGA 3413
3348 GAAGAAGTCAACCCCGAGCTTGCAGGAACCAAGTGGAGGCTGAAGTCAATTGTCCACTCTGA 3407
3414 CTTGGTGCATCCAAAGATGTCTGACCTCCACCTCCAGCACTGGAGCCAGAGGAACC 3473
3408 CTTTAGTGCACTTAACGGGAACCTTGACCTTCATCTCCAGACCTTGGAGCTTGGAGGCC 3467
3474 ACTGACTGCAGAGACTTGTGCTTCCAGCTCTGAGAGCTGTGATCTGTCTCAAGGAGCAGA 3533
3468 CTTGCTCCAGAGGCTCTGATCTCATCTCGGTGTTGGGATCCAGGCAAGGGGCGAGC 3527
3534 CTGCTGGGCAAGGAGCTGGAGGCTGCAACCAACCAAGTGGGCGCAGAGAGGCTCAC 3593
3528 CTGCTGGACAGAGGAGTGGAGGCTGTGAGCTGGCAAGCCCGGCGCCAGACAGACTTAC 3587
3594 CTGCTTCCAGAAAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3653
3588 CTGCTTGCAGAGGAGCAGTGTCTTCTGCTCTTACCGGACCTCCAGCAGCGGCGAGGT 3647
3654 TATAGAGGAGGCGCTTGGGAAAGCTGCGAGCCCAAGGCTGCTGCTGCTTCTTAAAGTCTAG 3713
3648 GCTAGAGGAGAGCCCTTGGAGATAGTGCAGTCAATCTCCCTGCTCTTAGGAGCCAG 3707
3714 CCCAAGCTTCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3747
3708 CCCAGGCTGCGCAGATCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3741
```

RESULT 5
AB211219
ID AB211219 standard; cDNA; 3450 BP.
XX AC AB211219;
XX DT 20-JAN-2003 (first entry)
XX Human polynucleotide SEQ ID NO 101.
DE Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nontoxic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.
XX OS Homo sapiens.
XX WO200270539-A2.
XX PD 12-SEP-2002.
XX PF 05-MAR-2002; 2002WO-US05095.
XX PR 05-MAR-2001; 2001US-0799451.
XX PA (HYSB-) HYSBQ INC.
XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX DR WPI; 2002-759812/82.
XX PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
PT platelet or coagulation disorders -

PS Claim 1; SEQ ID NO 101; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a

CC nucleotide sequence selected from any of 948 sequences

CC (ABZ1119-ABZ12066) or their mature protein coding portion, active domain

CC coding protein or complementary sequences. The polynucleotides are useful

CC for identifying expressed genes or for physical mapping of human genome.

CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular

CC weight markers, as a food supplement, for generating antibodies, in

CC medical imaging, screening and diagnostic assays and for treating

CC cell-proliferative disorders (cancer), neurodegenerative diseases

CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple

CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid

CC disorders, platelet or coagulation disorders, wound, burns, incision,

CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,

CC parasitic), arthritis, etc.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 3450 BP; 681 A; 1050 C; 1067 G; 652 T; 0 other;

Query Match 34.3%; Score 2118.8; DB 24; Length 3450;

Best Local Similarity 80.8%; Pred. No. 0;

Mismatches 2599; Conservative 0; Mismatches 567; Indels 50; Gaps 9;

991 CTTGCTGCCCGCCAGCCATCTCGAGGACCCGAGCGCTCTCGCGAGCGGGCCAGCAC 1050

DB 184 CTAGCGGCTCCCGCCATCACTACGCGGCCCGAGCGCTGTGCGGAGCGGGCGAGCA 243

QY 1051 GCGCGCTTCGTGCGCGCGCTCCGGGAGCCAGCGCCGCGCTGCACTGGCTGCAAC 1110

DB 244 GCGCGCTTCGTGCGCGCGCTCCGGGAGCCAGCGCTGTGCGGAGCGGGCGAGCA 303

QY 1111 GGGATCCCGTTGGGACCCCAATGGGCGCGTCAAGGTGCGAGGCGGTGCGCGACCTTGGTC 1170

DB 304 GGGGCGCGCTGCGGCGCCAAAGCGCGCTCAAGGTCCAGGCGCGGTGCGCGACCTTGGTC 363

QY 1171 ATCATCTAGATCGGCTGCGAGGACGCTGGCTACTACAGTGCCTAGAGAAACAGCGCG 1230

DB 364 ATCACAGATCGGCTGCGAGGACGCGCGCTACTACAGTGCCTGCGTGTGAGAAACAGCGCG 423

QY 1231 GGAATCGCTGCGCGCTGCGCGCTGCGCGGTAGTGTGCGCGAGGGGCTGCCAGCGCC 1290

DB 424 GGAATCGCTGCGCGCTGCGCGCTGCGCGGTAGTGTGCGCGAGGGGCTGCCAGCGCC 483

QY 1291 CCAGCTCGGCTCACAGCCAGCGCTGAGCAGCTCTCTGTGCTGTGGCTCGGAGCGG 1350

DB 484 CCACGCGGCTCACTGTAGCCCACTGAGCAGCTCCGCTGTGTGTGGCTGAGAGCGG 543

QY 1351 CTTGAGTTGACAGGAGCAAAATCAATTTGGCTTCTCTTCACTACCAAAAGGCAAGGGA 1410

DB 544 CCGGAGATGACAGCGAGCAGATCATCGGCTTCTCTCACTACCAAGGCAAGCGGCG 603

QY 1411 GTGCAATGTGGAGTACCAAGTTTGCAGTAAACATGACACACAGAGCTGCAAGTTGG 1470

DB 604 ATGGACAATGTGAATCAAGTTTGCAGTAAACATGACACACAGAGCTACAGAGTTGG 663

QY 1471 GACTTGAACCCCAACAGGATTAAGTTTCACTGTGTGGTGCCTACTCCAGCTGGGCGCC 1530

DB 664 GACTTGAACCCCAACAGGATTAAGTTTCACTGTGTGGTGCCTACTCCAGCTGGGAGCC 723

QY 1531 AGCGCAACCTTCAGCGCCAGCGCTGGTGACATGAGCGATGTCCCGAGCGAGCACCC 1590

DB 724 AGCGCAACCTTCAGCGCCAGCGCTGGTGACATGAGCGATGTCCCGAGCGAGCACCC 783

QY 1591 CAGCTTACCTTGTTCAGCGCCAGCGCTGGTGACATGAGCGATGTCCCGAGCGAGCACCC 1650

DB 784 CAGCTTCTCCCTGTTCAGCGCCAGCGCTGGTGACATGAGCGATGTCCCGAGCGAGCACCC 843

QY 1651 TCCAGCTGTAGCAATGACAGGTTGCTGAAGTACAGATACAGTGTGGGGAAGGAA 1710

DB 844 CCGAGCTGTAGCAATGAGGAGGTTGGTGAAGTACAGATACAGTGTGGGGAAGGAA 903

QY 1711 GATCAGGTTTTCTCCACCGAGGTGCTCGAAATGAGACACAACTTACGTTAACTCACCTT 1770

DB 904 GATCAGATTTTTCTCTACTGAGGTGCGAGAAATGAGACACAGCTTATGCTGAATCGCTT 963

QY 1771 CAGCCAAAACAAAGTGTACCGAGTCCGGATTTTCAGCTGGCAGCTGGCGCTATGGAGTC 1830

DB 964 CAGCCAAAACAAAGTGTATCGAGTACGGATTTTCGGCTGGTACAGCAGCGGCTTCGGGGCC 1023

QY 1831 CTTTCTCAGTGTGATGACAGCAGGACACTGTGTGTGCAACACAGAGCATGTTCCCTTT 1890

DB 1024 CCGTCCAGTGTGATGATCATCAGAGGACGCGCAGTATGCAACACAGAGCATGTCCTTTT 1083

QY 1891 GCGCTTCGAGATTCGAAGTGTAGGAGCAAGATGAGTCCCTGTGTGTGTCATGGCAGCG 1950

DB 1084 GCGCTTCGAGATTCGAAGTGTAGGAGCAAGATGAGTCCCTGTGTGTGTCATGGCAGCA 1143

QY 1951 CCGCTTCACCCCAACCCAGATCTCTGAGTACAAACTCTACTGGGAGAGGTGGGAAACAG 2010

DB 1144 CCGCTTCACCCCAACCCAGATCTCTGCTACAAACTATATTGGCGGAGGTGGGGCTGAG 1203

QY 2011 GAGGAGGAGATGTGTGACCGCGCCCGAGGGGTGCTGGAGATCAAGCTTGGGAGCTGGGG 2070

DB 1204 GAGGAGGCAATGTGGGATCGCGCTGCGAGGGGCGCTGGAGACAGGCTTTGGGATGTGGG 1263

QY 2071 CCGGTGCGGCTCAAGAGAAAGTGAAGCAGTATGAAGTATGAGTCTACCCAGCTAGTCTT 2130

DB 1264 CCGGTGCGGCTCAAGAGAAAGTGAAGCAGTATGAAGTATGAGTCTAGTCTTCCCTGG 1323

QY 2131 CCGTACGAGTGAAGCTCGTAGCTTTCAACAAACAGAGGAGCGGTACGCTGTGTGTGG 2190

DB 1324 CTTGACGAGTGAAGCTCGTAGCTTTCAACAAACATGAGATGGCTATGACGAGTGTGG 1383

QY 2191 AAGGGCAAGACGGAGAGCGCGCCAGCGCTATCCAGAGGGGGCCAGCGCTG 2250

DB 1384 AAGGGCAAGACGGAGAGCGCGCCAGCGCTATCCAGAGGGGGCCAGCGCTG 1443

QY 2251 CCGTTCGCCATGTCACGAGCAGATCAACAGCTCCACTTCCATTTGGCTTCGGTGGAG 2310

DB 1444 CCGTTCGCCATGTCACGAGCAGATCAACAGCTCCACTTCCATTTGGCTTCGGTGGAA 1503

QY 2311 AAGCCAGATTTTACCACTGTCAAGATTTGCAACTACACTGTAGCTTCGCGCCCTGGGG 2370

DB 1504 AAGCCAGATTTTACCACTGTCAAGATTTGCAACTACACTGTAGCTTCGCGCCCTGGGG 1563

QY 2371 CTTAGAAATGCTTCCCTGTGTACTATATACAGCTCTGGAGAGACATTTCTCATTTGGC 2430

DB 1564 CTTAGAAATGCTTCCCTGTGTACTATATACAGCTCTGGAGAGACATTTCTCATTTGGC 1623

QY 2431 GCGCTGAACCACTTACCAAGTACGAGTTTGGGTACAGTCCCGAGGTGGATATGGAT 2490

DB 1624 GCGCTGAACCACTTACCAAGTACGAGTTTGGGTACAGTCCCGAGGTGGATATGGAT 1683

QY 2491 GGGCGCTTTGGGCTTCGCTGTAGAACCTCCAGCTCCAGCGCTTCAGCGCTTCACACCTCCT 2550

DB 1684 GGGCGCTTTGGGCTTCGCTGTGTGTGGAGCGCTCCAGCGCTTCAGCGCGCTTCACACCGCC 1743

QY 2551 TCTGACCTGCGCTGAGCGCGCTGACCACTCCAGCGTTGCGTTACCTGCTGCTGCTGCTGCT 2610

DB 1744 TCGACCTGCGAGTGTGAGCGCGCTGACCACTCCAGCGTTGCGCTGCTGCTGCTGCTGCTGCT 1803

QY 2611 ACGGAGCCCAATGTTGAGATTTGAGATTTCTAATTTCTTACAGCAACACCAACCGAG 2670

DB 1804 ACGAGGCCCAACGGGGAGATCGTGGAGTATCTGATCTCTGTACAGCAGCAACCAACCGAG 1863

QY 2671 CCGGAACACAGTGGAGCACTGTCTCACACAGAGGGGAAACATCTTCAAGTGCAGAGGTCCAT 2730

DB 1864 CTTGAGCAACAGTGGAGCACTGTCTCACACAGAGGGGAAACATCTTCAAGTGCAGAGGTCCAT 1923

QY 2731 GCGCTTAGAGTGACACTCGGTATTTCTTCAAGATGGGAGCGCCGACAGAGGTGGGGCCT 2790

DB 1924 GCGCTTAGAGTGACACTCGGTATTTCTTCAAGATGGGAGCGCGGACAGAGGTGGGACCT 1983

QY 2791 GGGCCCTTTTCCCGCTTCAGAGTGTGATTACTCTGCAAGAGACATTTCTCAGACTCTCTTG 2850
|||||
Db 1984 GGGCCCTTTTCCCGCTTCAGAGTGTGATTCACTCCAGAGAGAGCTGTCTCAGACTCTCTTG 2043
|||||
QY 2851 GATGTGCAAGCGCTTCACCGGATCATCTGTGGTGTCTGCTGGGCTTCTCTCTGCTCTCTG 2910
|||||
Db 2044 GACATGCACTCAGTTCACCGGATCATCTGTGGTGTCTGCTGGGCTTCTCTCTGCTCTCTG 2103
|||||
QY 2911 GCTGTGATGTGTCTGCTGCTACGAAAGCTCCACAGGGAAGCCCTTCCCGGATGTGCC 2970
|||||
Db 2104 GCTGTGATGTGTCTGCTGCTGCGCGCAGCGCCACAGGGAATCCCTCCAGGCGTGTGCC 2163
|||||
QY 2971 TCCTCAGGACCCAGGAAACCCAGCGCTCTACACAGAGCTCGGCTGGGCTGCCAGT 3030
|||||
Db 2164 TCCACCGCACCCCGGATCTCCGCGCTGTACTCCAGAGCTGTGGCTTGGCCCGCCAGC 2223
|||||
QY 3031 GTCCCTGTGCTCAGTGTGGAGTCCCTCGTGTGATCTCTGCTCCCGAGGATGTGTGCCCA 3090
|||||
Db 2224 CCCCAGCTGCCATGATTTGGAGTCCCTTGTGTGACCCCATCCCGAGGACTGTGTCCCG 2283
|||||
QY 3091 CCACCTCAGATGTGGAAGACAGAGCTGAAGTACACAGCTTATGGGTGGCAGTGTTCAC 3150
|||||
Db 2284 CCACCTCAGAGTGGAGACAGAGGCTGAAGTGCACAGCTTATGGGTGGCGTGTTCCT 2343
|||||
QY 3151 GATTGCCGGGCGCACTCCAGAGAAAGATCTCTGGGCTCAGGACAGGGGAGCAACTGG 3210
|||||
Db 2344 GAAGCCCGAGTCACTCCAAAGAAAGATCTCTGGGCTCAACCAAGCGGGCTGAGCTGG 2403
|||||
QY 3211 GCAGCTCTCTGGCAGGCTGTGAGCTGCCCGCAGGCTAGTGTGTCAGAGCGGCTCTGACC 3270
|||||
Db 2404 GCTGGTCTCTGGCAGGCTGTGAGCTGCCCGCAGG---CAGGCCCGCGCGCTGTGACC 2460
|||||
QY 3271 CGTGTCTGTCTCTCCAGCGGAAACCGGGCAGACACTGTCTGTCTGCAAGCCCTGTGTAT 3330
|||||
Db 2461 CGGGCCCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
|||||
QY 3331 GACGGCATAAAGAGCAACCGGAGAAAGAGCGTGTCCCGAGCTGTGAGGAATCAGGTGAA 3390
|||||
Db 2521 GACGCCATAAAGGCAATGGGAGGAAGAGTCAACCCCGAGCTGTGAGGAACAGGTGGAG 2580
|||||
QY 3391 GCTGAGGTCTATGTCTCAGCTTCGAGTTCGAGTTCAGAGGATGTCTGAGCTCCAGCTC 3450
|||||
Db 2581 GCTGAGGTCTATGTCTCAGCTTCGAGTTCAGAGGATGTCTGAGCTCCAGCTC 2640
|||||
QY 3451 CAAGACCTGGAGCAGAGGAAACCACTGACTCAGAGACTCTGCTTCCACGCTCTGGAGCT 3510
|||||
Db 2641 CAAGACCTGGAGCTGAGGACCCCTGCTTCCAGAGGCTCTGATCTCATCTCGGTGT 2700
|||||
QY 3511 GTGATCTGTCTCAAGAGACAGACTGTGCTGGGCAAGAGCTGGAGGCTGCAACCAACA 3570
|||||
Db 2701 GGGGATCCAGGCGAGGGGAGCTGTGCTGACAGGAGTGTGGAGGCTGTGAGCTGGCA 2760
|||||
QY 3571 ACCAGTGGGCGAGAGGCTCACTGTCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 3630
|||||
Db 2761 GCGCGCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2820
|||||
QY 3631 TCAGACCTCCAGCGCAGCACTGTCTATAGAGGAGGCGCTGGGAAAGCTGCGAGCCAAA 3690
|||||
Db 2821 CCGGACCTCCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2880
|||||
QY 3691 GCTGTGTCTCTTAACAGTCAAGCCCTTCCAGGGGCGCTGTCTCTCTCTCTCTCTCTCT 3750
|||||
Db 2881 TCCCGCTGCTCTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940
|||||
QY 3751 GTCCCTGAGCAGA-----AGGAGATATGGCTCAGGAA----- 3784
|||||
Db 2941 TAGCTCTTCCAGAGGATGTGGTGGGCGAGGAGGATATGATACATAGGATGCGATA 3000
|||||
QY 3785 ---CATGCCATGATGCTACACATGTGTACTAGAGATATCCATAGTCTCTGGAGCC 3841
|||||
Db 3001 CCTGTGGCGGTGTATGCTCAATGTGTGCTGTAGATATCATCAAGCCCTTTGGAGCT 3060
|||||
QY 3842 TCTTAGGCTCTTTGGCTGGGGTGGGAGAA-----CTTTACTCTCTCTCATATTTCTG 3895
|||||

Db 3061 TCCTAAGTTGCTTTGGCTGAGGGAGAGAAACATGGATTATTCTACTCTCCCGCATACTC 3120
|||||
QY 3896 CATCACATACAGGAGGAGCTTGAGACACAGCTCTGTGTATGGACACGCTGTGAAGTCTGTG 3955
|||||
Db 3121 TTTGTGATACATGTGACATGTGAAGACATACGACATAGCTACATGTGATGTGCAC 3180
|||||
QY 3956 TGT 4010
|||||
Db 3181 ATGTGTGAAGTGTGATGT 4240
|||||
QY 4011 ACTCACGTGTGCTGT 4068
|||||
Db 3241 GGT 3300
|||||
QY 4069 GCTCTTTGAGCTTTTGT 4124
|||||
Db 3301 AGCCCTGTGGCGCCCACTTCCATTTGCGCCAGGCTTTTATTACACACTCTGAGAGTGTCT 3360
|||||
QY 4125 CCAATGCTGTCTTTGACAAAGACAGTTCCTCCAGGCT 4159
|||||
Db 3361 CCAATGCTGTCTTTGACAAAGACAGTTCCTCCAGGCT 3396
|||||
RESULT 6
AAI67203
ID AAI67203 standard; DNA; 2223 BP.
XX
AC AAI67203;
XX
DT 11-FEB-2002 (first entry)
XX
DE Nucleotide sequence of GSK gene Id 27142.
XX
KW Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;
KW antidiabetic; antidiabetic; anorectic; antianorectic; antiasthmatic;
KW antidiabetic; antidiabetic; anorectic; antianorectic; antiasthmatic;
XX cytosolic; cerebroprotective; vasotropic; human; db.
XX Homo sapiens.
XX W0200172961-A2.
XX
XX 04-OCT-2001.
XX
XX 22-MAR-2001; 2001WO-US09226.
XX
XX 24-MAR-2000; 2000US-192158P.
XX 28-MAR-2000; 2000US-192668P.
XX 27-APR-2000; 2000US-200166P.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RP, Xiang Z, Kabnick KS;
PI Lai Y;
PI
XX WPI; 2001-639223/73.
XX P-PSDB; AAG65913.
XX
XX Isolated polypeptides, which may be peptide hormones, which are
XX identified by high throughput genome-based biology which identifies
XX genes and gene products as therapeutic targets for treatment of
XX diseases such as diabetes and cancer -
XX
XX Claim 2; Page 54-55; 99pp; English.
XX
XX The invention provides polypeptides (AAG65986-65918) which may be peptide
XX hormones (including insulin, growth hormones, chemokines, cytokines,
XX neuropeptides, integrins, kallikreins, lamin, melanins, natruietic
XX hormones, neurotrophin, pituitary hormones, pleiotrophins, prostaglandins,
XX secretogranins, selectins, thromboglobulins, thymosins) identified by
XX high throughput genome-based biology and polynucleotides (AAI67176-67208)
XX encoding them. The polypeptides can be expressed by standard recombinant
XX

RESULT 8

AAC79893

ID AAC79893 standard; cDNA; 680 BP.

XX

AC AAC79893;

XX

DT 09-FEB-2001 (first entry)

XX

DE Human secreted protein encoding cDNA for gene 45.

XX

XX

XX

KW Human; secreted protein; cytosolic; antiarthritic; antiasthmatic;

XX

KW immunosuppressive; antiarteriosclerotic; antiinflammatory; nootropic;

XX

KW neuroprotective; antidiabetic; tranquiliser; vulnerary; antibacterial;

XX

KW antipariatic; antiarrhythmic; antiasthmatic; antirheumatic; cardiant; anti-HIV;

XX

KW autoimmune disorder; allergic condition; cardiovascular disorder;

XX

KW cancer; neurological disease; tissue repair; ss.

XX

OS Homo sapiens.

XX

XX

XX

PN W0200055176-A2.

XX

XX

XX

PD 21-SEP-2000.

XX

PF 09-MAR-2000; 2000WO-US06057.

XX

XX

XX

PR 12-MAR-1999; 99US-0124142.

XX

PR

XX

PR 11-JUN-1999; 99US-0138597.

XX

PR

XX

PR 03-DEC-1999; 99US-0168666.

XX

XX

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

XX

XX

PI Rosen CA, Ruben SM, Komatsoulis G;

XX

XX

XX

DR WPI; 2000-638176/61.

XX

XX

XX

DR P-PSDB; AAB44874.

XX

XX

XX

DR Novel 49 human secreted proteins useful for diagnosis, prevention and

XX

PT treatment of disorders including neurological, cell proliferative,

XX

PT cardiovascular, and autoimmune/inflammatory disorders and microbial

XX

PT infections -

XX

XX

XX

PS Claim 1a; Page 358; 405pp; English.

XX

XX

XX

CC This invention describes a novel isolated polypeptide (I) comprising an

XX

CC amino acid sequence at least 95 % identical to a polypeptide sequence

XX

CC selected from 49 polypeptides encoded by polynucleotide sequences

XX

CC included in American Type Culture Collection (ATCC) deposit number

XX

CC 203917, defined in the specification. The products of the invention have

XX

CC cytosolic, antiarthritic, antiasthmatic, immunosuppressive, nootropic;

XX

CC antiarteriosclerotic, antiinflammatory, neuroprotective, antidiabetic,

XX

CC tranquiliser, vulnerary, antibacterial, antipariatic, antiarrhythmic,

XX

CC antirheumatic, cardiant and anti-HIV activity. (I) or a nucleic acid (II)

XX

CC encoding (I) is useful for preventing, treating or ameliorating a medical

XX

CC condition and for diagnosing a pathological condition or susceptibility

XX

CC to the condition. (I) is useful for identifying a binding partner which

XX

CC affects the activity of the polypeptide and for identifying an activity

XX

CC in a biological sample. (I), (II) or an antibody (IV) specific to (I) is

XX

CC also useful for treating or preventing a disease, disorder or condition

XX

CC associated with aberrant expression of (I). Diseases treated or diagnosed

XX

CC include immune disorders such as autoimmune diseases, blood protein

XX

CC disorders, anemia, allergic reactions and conditions such as asthma,

XX

CC organ rejection or graft-versus-host disease, inflammation, hyper

XX

CC proliferative disorders, cardiovascular disorders such as arterioarterial

XX

CC fistula, arrhythmias, arteriosclerosis, coronary thrombosis, organ

XX

CC regeneration, cancer, neovascular glaucoma, diabetic retinopathy,

XX

CC rheumatoid arthritis, psoriasis, diseases associated with increased

XX

CC apoptosis that include acquired immunodeficiency syndrome (AIDS),

XX

CC neurological diseases such as Parkinson's disease, viral, bacterial,

XX

CC fungal or parasitic diseases. They are also used to repair, replace or

XX

CC protect tissue damage by congenital defects, to treat trauma, in surgery,

XX

CC including cosmetic plastic surgery, to treat fibrosis, reperfusion injury

XX

CC or systemic cytokine damage, to stimulate chondrocyte growth, to prevent

XX

CC skin aging due to sunburn, to change a mammal's mental state or physical

CC state by influencing biorhythms, cardiac rhythms, depression, memory,
CC stress and for accelerating wound healing. (I), (II) and/or their agonist
CC or antagonist are useful as food additives or preservatives to increase
CC or decrease storage capabilities, fat content, lipid, protein,
CC carbohydrate, vitamin, mineral or other nutritional components. (I) is
CC useful for screening therapeutic compounds. (II) is useful in forensic
CC biology for detecting DNA sequences and as diagnostic probes for
CC detecting the presence of specific mRNA in a particular cell type.

XX
SQ Sequence 680 BP; 251 A; 97 C; 100 G; 232 T; 0 other;

Query Match 5.5%; Score 337.2; DB 21; Length 680;
Best Local Similarity 79.9%; Pred. No. 5.5e-77;
Matches 473; Conservative 0; Mismatches 108; Indels 11; Gaps 6;

QY 5588 GAATGAAATAACCGTTTAAACAGTAGGCTCTTAGCATCACACATAGCTCTCTCAT 5647
DB 11 GAATTAATGGCTGTGTTTAAACAGTAGGCTCTTAGCATTAATACCATAGCTCTCTCAT 70

QY 5648 GTTCTTGTGTTTAAACAGCACTTGAAGGTTCTGGGTTTAAATTAATAGTGCATGAGACA 5707
DB 71 GTTCTTGTGTTTAAACAGCACT--GAGGTTCTGGTGTAAATTAATAGTGCATGAGACA 128

QY 5708 ATTATTAACCCATTAGGCTGGGTGGAATTTG-TTCTCAAAAGCAAAATAGTAATAATC 5766
DB 129 ATTATTAACCCATTAGGCTGGGTGGAATTTGTTCTCAAAAGCAAAATAGTAATAATC 188

QY 5767 TGGTATCTGCCCTATTAACCTCACAGTTGTATAAGAAAGTAGGCAGAACCTCCTAGCATTAAT 5826
DB 189 TGGTATCTGCCCTATTAACCTCACAGTTGTATAAGAAAGTAGGCCTTTCTCCTAGCATTAAT 248

QY 5827 ATGATTTGGGGTCTCGAGTAACTGGGGAGTGTGTAGCTTTTGTAGTCTTTGTAGCACCAGGTCT 5886
DB 249 ATGATTTGGGGTCTCGGGTAATTTGGAGTGTGTAGGTTGTGTCTTTGTAGCAGTATTTT 308

QY 5887 TAATTAGGAA-AGTCTGTGGCCTTTTACAGGCAATAGTCCCTTTGTCTTTGTGCCATGGA 5945
DB 309 ATTAGAAAAGAACTATTGGCCCTTTTACAGGGTATTAATCCCTTTGTCTACCTACCATGGA 368

QY 5946 TGCCTTAAGTCTTTGGAGTCTCATTTAAGAA-TTCTTTTCTCGAAGCATGACAAGTGT 6004
DB 369 TGCCTTAAGTCTTTGTAGTCTCAATTAAGAAATCTTCTTTTGTAGCATGACAAGTGT 428

QY 6005 ATCGCAATCTTACATGCTCACTCGTTTACCTGGCTTAGTGTGTGCTGGGTATTTAAAT 6064
DB 429 -----AATCAGTACTTGTCTCATTTATTGTCTGTATTAGTGTACTTATTTAAT 483

QY 6065 GCATTTTCAGCATCATGCTTCTCTTCAAAATATGATATTTCTTTATTGTACACTAAG 6124
DB 484 TATCCTTCCAGCGGTTTTTTTTTCTCCTTACAAATATGATATACTCTTTAGTGTAAAGCTAAG 543

QY 6125 GTGTTGA-TCATGTATCTGCTCCCTGTAAGAAATTAATAACTATTTTCCAGA 6175
DB 544 GCATTTGATCTATGATCTGCTCTTATTAATGAATTAATAACTATTTTCCAGA 595

RESULT 9

ABX71437

ID ABX71437 standard; cDNA; 2976 BP.

XX

AC ABX71437;

XX

DT 06-MAR-2003 (first entry)

XX

DE Human cDNA encoding a novel tyrosine phosphatase, NHP6.

XX

KW Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;

XX

KW Ig super family protein; gene therapy; NHP; novel human protein.

XX

OS Homo sapiens.

XX

XX

XX

PN US6465632-B1.

XX

XX

PD	15-OCT-2002.
XX	
PF	08-JUN-2001; 2001US-0877730.
PP	
PR	09-JUN-2000; 2000US-210607P.
PS	(LEXI-) LEXICON GENETICS INC.
PA	
XX	
PI	Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;
PI	Zambrowicz B, Sands AT;
DR	WPI; 2003-147071/14.
DR	P-PSDB; ABU5420S.
XX	
PT	Novel isolated nucleic acid which encodes a novel human protein that
PT	shares sequence similarity with animal phosphatases, that is useful for
PT	generating antibodies, and as reagents in diagnostic assays -
XX	
PS	Disclosure; Column 49-52; 58pp; English.
XX	
CC	The invention relates to an isolated nucleic acid molecule comprising a
CC	nucleotide sequence encoding a novel human protein (NHP) that shares
CC	sequence similarity with animal phosphatases (in particular tyrosine
CC	phosphatase and is a member of the Ig superfamily). The disclosed NHP
CC	polynucleotide sequences (from genes located on either chromosome 15
CC	or chromosome 3) are useful for identifying coding sequences and in the
CC	identification of biologically relevant splice junctions. The NHP
CC	polynucleotide sequences are useful in gene therapy, for detecting or
CC	mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
CC	disease, for screening drugs effective in treatment of symptomatic or
CC	phenotypic manifestations of perturbing the normal function of NH.
CC	Complementary sequences of the NHP polynucleotides used in conjunction
CC	with PCR to screen libraries, isolate clones and prepare cloning and
CC	sequencing templates. Labeled NHP nucleotide probes can be used to screen
CC	a human genomic library which is helpful for identifying polymorphisms,
CC	determining the genomic structure of a given locus/allele and defining
CC	diagnostic tests. The probe sequences also have use in defining and
CC	monitoring both drug action and toxicity. Oligonucleotides
CC	complementary to NHPs may encode or act as NHP antisense molecules, or
CC	may be used as part of ribozyme and/or triple helix sequences.
CC	The present sequence encodes an NHP.
XX	
SQ	Sequence 2976 BP; 844 A; 697 C; 653 G; 781 T; 1 other;
	Query Match 5.1%; Score 316.6; DB 25; Length 2976;
	Best Local Similarity 47.4%; Pred No. 3.1e-71;
	Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6
QY	300 AGACGCGATGATGAGGNAGCTCTTAGGATCTGGAAAGTCACTGAGGCAGCATATTTCGT 359
Db	45 ATACATCAGTGAGGTGGGAAGCAGCGCAGGAGAGCAGTCCGATGAAGGATTTATTCAGTG 104
QY	360 TCTGGCCCCAAGCCGGCTAGAGATGTGTGCCAGCAGGTTGCTGTGTGCTCAAGCTTGGCAC 419
Db	105 CTTGCGCAATGAACAATAATGAGGCCATCTTAGTCAAAGAAGTCACTTGGCTTATCAAC 164
QY	420 ACTCGAAGACTTCTCTCTGCAACCCTGAGTCCAGATTCTGGAGGAGAACGGGACAGCAG 479
Db	165 TAATTTCTGCATTGGAATTCAGCCAAATTTCACTGAGGTCACGAAGTGGAGTTGCTGTG 224
QY	480 CTTTGAATGCCACACCAAGGCGCTTCCAGCCCCCATCAATTACTTGGGAAAAGGACCAAGT 539
Db	225 ATTTCGATGCAAGATTTTCATCCCAACCTCTCTGCAGTCATAACATGGAGTTCAATCGGC 284
QY	540 GACCGTGCTGAG---GAGCCCCGGCTCATCACTCTTCCCAAGTGGCTCTCTCCAGATCCT 596
Db	285 AACCTCTACCTATGACTACTGACAGGATAACTCGCCCTACCAACAGGAGTATTGCAGATCTA 344
QY	597 AGATGTCAGGACAGTGTAGCAGCTCTCTACCGTGTGGTGCCCAACCAATTCAGCCCGCCA 656
Db	345 TGATGTGAGCCAAAGGGATTCGGAANAATTCGTTGTATTGTGTCACATGTAGGCCACCG 404
QY	657 ACGATTCCAGCCAGAGGCGCTCGCTCACTGTGGCCCTCTCAGAGGGGTCTTTGGAGGGCTACCAG 716

Db	1473	CCCGGGGACACGCGATGAGTACCTTTTGGAGGCCGTGAACCTGACAGTGTCTACTCTGGT	1533
Qy	1794	CCGGATTTCAGCTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGGATGACGACAG	1853
Db	1533	TGCGATTACTGCTGCCACAGAGTGGGGCTGGGAGTCAATCAGTATGGACTTCAATAG	1592
Qy	1854	GACACCTGGTGTGCAAAACAGAGCGAATGTTCCCTTTGCCCCCTGCGAGAAATTGAAGTGTAG	1913
Db	1593	GACGCCAAAGCTACAAAGCGTGAAAGCCCTTAAGTCTCCAGAGTTGCAATTTGAGGCGCTCT	1652
Qy	1914	GGCAAGATGGAGTCCCTGGTGTGTCAATGGCAGCGCCGCCCTCACCCACCCAGATCTC	1973
Db	1653	GAACTGTACCACTTTCTGTGAGGTGGCAGCAGATGTAGAGGACACAGCTCTCTATTCA	1712
Qy	1974	TGGATACAAACTCTACTGGGGAGAGGTGGGAAACAGAGGAGGAGCAGATGGTACCGCCC	2033
Db	1713	GGGCTACAGCTGTACTACAAGGAGNAGGGCAGCAGGAGAA	1754
Qy	2034	CCCAGGGGTGTGGAGATCAAGCTTGGGACGTGGGCCCGTGGCGCTGAAGAAGAAAGT	2093
Db	1755	-----TGGGCCCAATTTCTTGATACCAAGGA	1781
Qy	2094	GAAGCAGATGAAGTGAACCTGAGTTAGTCCCTGGCAGGCCGTACGAGGTGAAGTCTCGTAGC	2153
Db	1782	CCTACTCTACTCTCTCAGTGGCTTAGACCCCAAGAGAAATATCATGTGAGACTCTCTGGC	1841
Qy	2154	TTTCAACAAACACGAGGACGGCTACGCTGTGTGTGAAGCGGCAAGACGGAAGGCGCC	2213
Db	1842	TTACAAACATATAGACGATGGCTATCAGGCAGATCAGACTGTCAAGC-----TCCAGGATG	1898
Qy	2214	CACGCCAGACCTGCCCTATCCAGAGGGGGCCACCGCTGCCTCTGCCCAATGTCACGCAGA	2273
Db	1899	CGTGTCTGTGGTATCGCATGTGCTCTCCACCAACACCCACCCTCTCTATGCGAA	1958
Qy	2274	GTCAACAGCTCCACTTCATTTGGCTTCGGTGGAGAGGACGAGACTTTACCACTGTCAA	2333
Db	1959	GGTAAACACCTCATCTTCCATCTCTCCGCACTGGAGGAGCCCTGCATTCACCCCTGCACA	2018
Qy	2334	GATTGTCAACTACACTGTACGTTTGGCCCTTGGGGCTCAGGAATGCTTCCCTGGTGCAC	2393
Db	2019	AATCATTTAACTACACATCCGCTGTAACTCTGTGTGGCTTGAGAAATGCTTCTTTGGTCT	2078
Qy	2394	CTACTATACAGCTCTGGAGAAGCAATTCTCATTTGGCGGCTGAACCAATTTTACCAAGTA	2453
Db	2079	GTACCTTCAAAACATCAGAAACTCATATGTTGTTTCAAGGCTTAGAACCAACCAAAATA	2138
Qy	2454	CGAGTTTGGGTACAGTCCACGAGTGGATATGATGGGGCCCTTTGGCTCCGTCGTAGA	2513
Db	2139	CGAATTTTGGCGTTTCGATATCATGTGGATCAGCTTTCCAGTCTTTGGAGCCCTGTAGTCTA	2198
Qy	2514	ACGCTCCACCTCGCTCAGCGGCTTCAACACCTCTCTTCTGACTCGCTGGCTGAGCCCCCT	2573
Db	2199	CCATTCTACTCTCCAGAGCACACGAGGCCCAACAGTTGGAGTAAAGTGACATTAAT	2258
Qy	2574	GACACCATCCACCGTTCCGTTTACACTGGTGTCCCCCAACGAGGCCCAATGGTGGAGTTGT	2633
Db	2259	AGAGGATGACACTGCGCTGTTTCTTGGAAACCCCTTGATGGCCCAAGACAGTTGTGAC	2318
Qy	2634	GGAGTATCTAATTTCTTACAGCAACAAACCAACCCAGCCCGAAACACAGTGGACACTGCT	2693
Db	2319	CCGCTATACTATCTTATATGATCTAGAGAGGCCCTGATTTGCAGGAGTGGCAGGTCTT	2378
Qy	2694	CACACAGAGGGAACATCTTCACTGAGAGGTCCATGGCTAGAGAGTGCACCTCGGTA	2753
Db	2379	ACACCGTGAAGGGGCAATAACCATGGCTTTGCTAGAAAACTTTGGTAGCAAGAAATGTGA	2438
Qy	2754	TTTCTTCAAGATGGGAGCCCGCACAGAGGTGGGGCCCTGGGCCCTTTTCCCGCTTGACAGGA	2813
Db	2439	CATTGTCAAGATATCTGCATCCATGAGTGGGAGAGGAGCCCTTTCAAATCTGTGGA	2498
Qy	2814	TGTGATTACTCTGCAAGAGACATTTCTCAGACTC	2846
Db	2499	-----GCTGGCAGTACTTCCAAAGGAAACCTCTGAATC	2531

RESULT 10	
ABX71432	
ID	ABX71432 standard; cDNA; 3210 BP.
XX	
AC	ABX71432;
XX	
DT	06-MAR-2003 (first entry)
XX	
DE	Human cDNA encoding a novel tyrosine phosphatase, NHP1.
XX	
KW	Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;
KW	Ig super family protein; gene therapy; NHP; novel human protein;
KW	SNP; single nucleotide polymorphism.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	variation
FT	76
FT	Location/Qualifiers
FT	/*tag= a
FT	/standard_name= "Single nucleotide polymorphism"
FT	706
FT	variation
FT	/*tag= b
FT	/standard_name= "Single nucleotide polymorphism"
XX	
FN	US6465632-B1.
XX	
PD	15-OCT-2002.
XX	
PF	08-JUN-2001; 2001US-0877730.
XX	
PR	09-JUN-2000; 2000US-210607P.
XX	
PA	(LEXI-) LEXICON GENETICS INC.
XX	
FI	Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;
FI	Zambrowicz B, Sands AT;
XX	
DR	WPI; 2003-147071/14.
XX	
XX	P-PSDB; ABUS4200.
FT	Novel isolated nucleic acid which encodes a novel human protein that
FT	shares sequence similarity with animal phosphatases, that is useful for
FT	generating antibodies, and as reagents in diagnostic assays
XX	
PS	Disclosure; Column 13-18; 58pp; English.
XX	
CC	The invention relates to an isolated nucleic acid molecule comprising a
CC	nucleotide sequence encoding a novel human protein (NHP) that shares
CC	sequence similarity with animal phosphatases (in particular tyrosine
CC	phosphatases and is a member of the Ig superfamily). The disclosed NHP
CC	polynucleotide sequences (from genes located on either chromosome 15
CC	or chromosome 3) are useful for identifying coding sequences and in the
CC	identification of biologically relevant splice junctions. The NHP
CC	polynucleotide sequences are useful in gene therapy, for detecting
CC	mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
CC	disease, for screening drugs effective in treatment of symptomatic or
CC	phenotypic manifestations of perturbing the normal function of NH.
CC	Complementary sequences of the NHP polynucleotides used in conjunction
CC	with PCR to screen libraries, isolate clones and prepare cloning and
CC	sequencing templates. Labeled NHP nucleotide probes can be used to screen
CC	a human genomic library which is helpful for identifying polymorphisms,
CC	determining the genomic structure of a given locus/allele and designing
CC	diagnostic tests. The probe sequences also have use in defining and
CC	monitoring both drug action and toxicity. Oligonucleotides
CC	complementary to NHPs may encode or act as NHP antisense molecules, or
CC	may be used as part of ribozyme and/or triple helix sequences.
CC	The present sequence encodes an NHP.
XX	
SQ	Sequence 3210 BP; 895 A; 761 C; 716 G; 836 T; 2 other;

Query Match

5.1%; Score 316.6; DB 25; Length 3210;

Best Local Similarity 47.4%; Pred. No. 3.2e-71; Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6;			
QY	300	AGACAGCATGATGAGGAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCCTG	359
Db	279	ATACATCAGTGAAGTGAAGCAGGAGGAGGAGCAGTCCGATGAAGGATTTATCAGTG	338
QY	360	TCTGGCCACAGCCGCTAGGAGTGGTGGCCAGCAGGTTGCTGGTCAAGCTTGGCAC	419
Db	339	CTTGGCAATGAACAAATATGAGCCATCTTAGTCAAAAAGCTCATCTTGGCTTATCAAC	398
QY	420	ACTCGAAGACTTCTCTCTGCAACCCGAGTCCAGATTGTGGAGGAAGCGGACAGCAG	479
Db	399	TAATTTCTGCAITGGAAGTCCAGCCAAATTTCCACTGAGGTCCACGAAGTGGAGTGGCTCG	458
QY	480	CTTTGAATGCCACACCAAGGGCTTCCAGCCCCCATCATTAATCTGGGAAAGGACAGGT	539
Db	459	ATTTCATGCAAGATTTCAATCCACCCCTCTGCACTCATACATGGGAGTTCAATCGGAC	518
QY	540	GACCGTGGCTGAG---GAGCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCCT	596
Db	519	AACTCTACCTATGACTATGACAGGATAACTGCCCTCACCAACAGGAGTATTGCAGATCTA	578
QY	597	AGATCTCAGGACAGTGAAGGCTCTACCGCTCGGTGGCCACCAATTCAGCCCGCA	656
Db	579	TGATGTCAAGCCAAAGGGAATCTGGAAATTTATCGTTGTATTGCTGCTGACCTGAGCCACCG	638
QY	657	ACGATTCAGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAG	716
Db	639	AGTAAAGTATGAGGGCTCGCTAACTGTGATTCAG-----CTAAGGAGTCAAAATC	692
QY	717	GGGGCAGGATGGTCAATTTGTGGCAGCCCGAGAGAACCAACGGTAGTGTCTGGACAGAA	776
Db	693	CTTCCACACACCARCAATATAGCAGGTCACAGAAATACCAACATCTCTTCATCAGAC	752
QY	777	TGATGATGAGTGGTGGCTCTGCTGACCCCAACCTCTTGTGTCTGGTCCGACA	836
Db	753	TGTAGTTTGGAAATGATGGCCACAGGAAATCCCAAGCCAAATCATTTCTGGAGCGGCT	812
QY	837	GGATGAAAGCCTAT---CTCCACGGATGTCATCGTCTGGCCGAGCAATCTACTCAT	893
Db	813	TGATCACAATCCATTGATGCTTTTAATCTCGGGTACTTGGAAATGGTAAATCTCATGAT	872
QY	894	CGCCAGCGCGAGCCTCGGCACTCTGGAGTCTATGCTCTGCCAGGCAACAAAGCCCTCAC	953
Db	873	ATCTGATGTCAGGCTACAAACATGCTGGAGTATATGTTGTGGGCCACTACCCCTGGCAC	932
QY	954	GGTGACTTGGCACTGGGCTGCTGAGTCCGAGTGTGCTGCCAGCCCATCTCGCA	1013
Db	933	ACGCAACTTTACAGTTGCTATGGCAACTTTAACTGTATTAGCTCTCTCTTCAITTTTGA	992
QY	1014	GGCACCAGGCGCTCTCGCGACGCGGCGCAGCACCGCGCTTCTGGTGGCGGGCGTC	1073
Db	993	ATGGCCAGAAAGTTTAAAGAGCCCTCGAGCTGGCACTGCTCGATTTGTGTGTCGAGCAGA	1052
QY	1074	CGGGAGGCCACGGCCCGGGCTGCACTGGCTGCAACAGCGGATCCCGTTGGACCCCAATGG	1133
Db	1053	AGGAATCCCTCTCCAGAGTGTATGTTGAAATATGGAGGAAGATACATTCGAATGG	1112
QY	1134	GGCGCTCAAGTGCAGGGCGGTGGCGAGTGTGTATCACTCAGATGGCGCTGCAGGA	1193
Db	1113	TAGAAATTAATATGTAACACAGT-----AAATTTGTAATTAACACAGATTAATTCCTGAAGA	1166
QY	1194	CGCTGGCTACTACAGTGGTAGCAAAACAGCGGGGAAGTGGCTGTGGCTGGCGCC	1253
Db	1167	TGATGCTATTTATCAGTGCATGGCTGAGAAATAGCCAAAGATCTATTTATCTAGAGCCAG	1226
QY	1254	CCTGGCGGTAGTGGTGGCGAGGGCTGCCAGCGCCCGGACTCGGGTCAACAGCCACGCC	1313
Db	1227	ACTGACTAGTGTGTCAGAGACAGACCCAGTGTCTCCCTATATGTACATGCTGTAAC	1286
QY	1314	GCTGAGCAGCTCTCTGTGGTGGCTGGAGCGGCTGTGATTTGCAACAGGACAAAT	1373

Db	1287	CATGTCAGGCTCAGCCATTTCTTTAGCCTGGGAGAGGCCACTTTATATTTAGACAAAGT	1346
QY	1374	CATTGGCTTCTCTCTTCACTACCAAAAGCGAAGGGGAGTGACAAATGTGGAGTACCAGTT	1433
Db	1347	CATTGGCTTCTCTCTTCACTACATGAAGCAGAAAGGTTTAAATTAATGAAGAGTATCAAGT	1406
QY	1434	TGCAGTAAACAAATGACACCAACAGAGCTGAGGTTGGGAGCCTGGAAACCAACACGAGATTA	1493
Db	1407	AGTCATCGGAAATGACACCACTCATTTATTTATTTGATGACTTAGAGCCTGCGACAAATTA	1466
QY	1494	TGAGTTCTACGTGGTGGCTACTCCAGCTGGGGGCGCAGCGAACTCCACGCCAGCCCT	1553
Db	1467	TACTTTCTACATTTAGCATATATGCCAATGGGAGCCAGCCAGATGTCTGACCATGTGAC	1526
QY	1554	GGTGATACACTGGAACGATGTCCCAGCGCAGCACCCAGCTTACCTTGTCTCAGCCCCAA	1613
Db	1527	ACAGATACCTCTAGAGGATGTTCCCTTGAGACCTCTCTGAAATTTAGTTTGACAAGTCGAAG	1586
QY	1614	CCCTCGGACATCAGGGTGGCATGGCTGCCCTCGCCCTCAGCCTGAGCAATGACAGAGT	1673
Db	1587	TCCCACTGATATCTCATCTCTGGCTGCCAATCCAGCCAAATATCGCGGGGCCAAGT	1646
QY	1674	GCTGAAGTACAAATAGAGTACGGTGGGGAGGAGATCAGGTTTCTCCACCGAGGT	1733
Db	1647	GGTCTGTATCGCTTGTCTTCCGCTAAGTACTGAGAAITTCAAATCCAAGTTCTGGAGCT	1706
QY	1734	GCCTGGAAATGAGACACAACTTACGTTAAACTCACTTCAGGCCAAACAAAGTTGACCGAGT	1793
Db	1707	CCCGGGGACCAACGATGAGTACCTTTTGGGAAGGCTGAAACCTGACAGTGTCTACCTGCT	1766
QY	1794	CCGATTTTACGCTGGCCTGGCTATGGATGCTCTTCTCAGTGGATGACGACACAG	1853
Db	1767	TGGATTTACTGCTGCCACAGAGTGGGGCTGGGAGAGTCATCAGTATGCACTTCACATAG	1826
QY	1854	GACACCTGGTGGCAACACACAGAGCCATGTTCCCTTTGGCCCTCGCAGAAITTTGAAGTGA	1913
Db	1827	GACGCCAAAGCTCAAGCGGTGAAGCCCTTAAGTCTCCAGAGTTGCAATTTGGAGCTCT	1886
QY	1914	GGCAAGATGAGTCCCTGCTGGTGTCTATGGCAGCGCCGCTTCACTCCACCCAGATCTC	1973
Db	1887	GAACTGTACCACTTCTGTGAGTGGCAGCAAGATGTAGAGGACACAGCTGTCTATTTCA	1946
QY	1974	TGGATACAACTTCTTCTGGGAGAGTGGGAAACAGAGAGGAGGAGGAGATGTTGACCGCC	2033
Db	1947	GGGCTACAAGTGTACTACAAAGGAAGAGGCGCAGCAGAGAA-----1988	1988
QY	2034	CCCAGGGGTGCTGGAGATCAAGCTTGGGACGTGGGGCCGCTGGGCTGGAAGAAAGT	2093
Db	1989	-----TGGGCCCAATTTTCTTGGATACCAAGGA2015	2015
QY	2094	GAAAGCATATGAACCTGACCCAGTTAGTCCCTGGCAGGCGGTACGAGGTGAAGCTCGTAGC	2153
Db	2016	CCTACTCTATCTCTAGTGGCTTAGACCCAGAGAAATATCATGTGAGACTCTGCGC	2075
QY	2154	TTTCAACAAACACAGGAGCCTACGCTGTGTGTGGAAAGGCGCAAGACGGAAGGCGCC	2213
Db	2076	TTACAAACAAACATAGACGATGGCTATCAGGCAGATCAGACTGTCTGACAC---TCCAGGATG	2132
QY	2214	CACGCCAGACTGCTATTCAGAGGGGGCCACCGCTGCTCTCTGCCCATGTCTCAGCAGAA	2273
Db	2133	CGTGTCTGCTGATGCGCATGGTCCCTCTCTCCACCCACCCACCCCATCTCTATGCGAA	2192
QY	2274	GTCAAAACAGCTTCACTTCTCATTTGGCTTGGTGGAAAGCGACACTTTTACACCTGTCAA	2333
Db	2193	GGCTAACACCTCATCTTCCATCTTCTGCACTGGAGAGGCGCTGATTTACCGCTGCAACA	2252
QY	2334	GATTTCAACTACATGCTGACGCTTGGCGCTTGGGGGCTCAGGAATGCTTCCCTGGTCAAC	2393
Db	2253	AATCATTAACCTACACCTCCGCTGTAAATCTCTTGGCTGGCTGAGAAATGCTTCTTTGGTTCT	2312
QY	2394	CTACTATACAGCTCTGGAGAAAGATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTA	2453
Db	2313	GTACCTTCAMACATCAGAAACTCATGTTGGTTCAAGGCTTAGAAGCCAAACACCAATA	2372

QY	1074	CGGGAGCCACGGCCCGCGCTGCATCTGGCTGCACGACGGGATCCGTTTGGACCCAAATGG	1133
DB	819	AGGAATCCCTCTCTCCCAAGATGTTCATGGTTGAAATAATGGAAGGAGATACATTCGAATGG	878
QY	1134	CGCGGTCAAGGTGCAGGCGCGTGGCGGAGCTTCGTTCATCTCAGATCGGCTGCAGGA	1193
DB	879	TAGAAATTAATATGTAACAAGT-----AAATTGGTAATTAACAGATTAATCTCGAAGA	932
QY	1194	CGCTGGCTATACACAGTGCCTAGCAGAAACACGCGCGGAACTCCCTGTGCGCTGCGCC	1253
DB	933	TGATGCTATTTATCAGTGCTAGCTGAGAAATAGCAAGGATCTATTTATCTAGAGCCAG	992
QY	1254	CTTGGCGGTAGTGTGCGGAGGGCTGCCAGCGCCCGGACTCGGGTCAACAGCCAGCC	1313
DB	993	ACTGACTGTGTAGTGTGTCAGGAAGACAGACCCAGTGTCTCCCTATTAATGTATGCTGGAAC	1052
QY	1314	GCTGAGCAGCTCTCTGTGCTGGTGGCTGGGAGCGGCTCGAGTTGCACAGCGAGCAAT	1373
DB	1053	CATGTCAAGCTCAGGCATCTTTTAGCTGGGAGGGCCACTTTATTAATTCAGACAAGT	1112
QY	1374	CATTGGCTTCTCTCTTCAACCAAGGCAAGGGAGTGACAAATGTGGAGTACCAAGT	1433
DB	1113	CATTGGCTTCTCTGTACACTACATGAAAGCAGAGGTTTAAATTAATGAAGTATCAAGT	1172
QY	1434	TGCAGTAAACAATGACACACAGAGCTGCAGTTTCGGGACCTGGAAACCAACACAGGATTA	1493
DB	1173	AGTCATCGGAAATGACACAACTCATATATATTATGATGACTTAGAGCTGCGCAGCAATTA	1232
QY	1494	TGAGTTCTACGTGTGCGCTTACTCCAGCTGGGGGCGAGCGCACTTCAGGCCGAGCCCT	1553
DB	1233	TACTTTCTACATGTGTAGCATATATGCCAATGGGAGCCAGGCAGATGTCTGACCAATGTGAC	1292
QY	1554	GGTGATACACTGCACAGATGTCCCAGCGCAGCACCCAGCTTACCTTGTCTCAGCCGCA	1613
DB	1293	ACAGAACTCTAGAGGATGTCCCTCTGAGACTCTCTGAAATGATTTGACAGTGTGAG	1352
QY	1614	CCCTTCGGACATCAGGGTGGCATGGCTGCCCTGCCCTCCAGCCTGAGCAATGACACAGT	1673
DB	1353	TCCCACTGATATCTCATCTCTGGCTGCCAATCCAGCCAAATATCGCGGGGCCAAGT	1412
QY	1674	GCTGAAGTACAAGATAGAGTAGGTTGGGGAGGAGAGATCAGGTTTTTCTCCACGAGGT	1733
DB	1413	GGTCTGTATCGCTGTCTTTCCGCTAAAGTACTGTGAGAAATTCAAATCAAGTTCTGGAGCT	1472
QY	1734	GCCTGGAAATGAGACACAACTTACGTTAACTCACTTTCAGCCAAACAAAGTGTACCGAGT	1793
DB	1473	CCCGGGACACGATGATGATACCTTTTGGAGGCGCTGAAACCTGACAGTGTCTACCTGGT	1532
QY	1794	CCGGATTTTTCAGCTGGCATGCGCTGGCTATGGAGTCCCTTCTCAGTGGATGAGCAGCAG	1853
DB	1533	TCGGATTACTCTGCCACAGATGGGGCTGGGAGAGTCACTAGTATGGACTTCACATAG	1592
QY	1854	GACACTGGTGTGCACACACAGAGCCATGTTCCCTTTGCCCCCTGCAGAAATGAAGTGTAG	1913
DB	1593	GACGCCAAAGACTACAAGCGTGAAGGCCCTTAAGTCTCCAGAGTTGCATTTGGAGCCCT	1652
QY	1914	GGCAAGATGGAGTCCCTTGGTGGTGTATGGCAGCCGCCCTTCACCCCAACCCAGATCTC	1973
DB	1653	GAACTGTACCAACCATTTCTGTGAGGTGGCAGCAAGATGTGAGGACACAGCTGCTATTCA	1712
QY	1974	TGGATACAAACTCTACTTGGGAGAGGTGGGAAACAGAGGAGGAGGCAGATGGTGCAGCC	2033
DB	1713	GGGCTACAAGTGTACTACAAGGAAGAGGGCAGCAGAGNA-----	1754
QY	2034	CCCAGGGGTCTGTGGAGATCAAGCTTGGGACGTGGGCGCGGTGGGCTGAAGAGAAAGT	2093
DB	1755	-----TGGGCCCATTTTCTTGGATACCAAGGA	1781
QY	2094	GAAGCAGTATGAACCTGACCCAGTTAGTCCCTGGCAGGCGGTACAGAGTGAAGCTCGTAGC	2153
DB	1782	CCTACTCTATACTCTCAGTGGCTTAGACCCCGAGGAAGAAATATCATGTGAGACTCTCGGC	1841

Qy	2154	TTTTCACAAACACGAGGAGCGGCTACGCTCTGTGTGAGAGGCGAAGCGAGGAAGCGCC	2213
Db	1842	TTACAAACACATAGACGATGGCTATACGGCAGATCAGACTGTCAGCAC---TCAGATG	1898
Qy	2214	CAGCCACAGACTGCCTATCCAGAGGGGGCCACCGCTGCCTCTCGCCCATGTCCACACGAGA	2273
Db	1899	CGTGTCTGTGCGTATCGCATGTGCTCCTCCACACACCCACCACTCTCTATGCGAA	1958
Qy	2274	GTCAAAACAGCTCCACATTCGATTTGGCTTCGGTGGAGAGGCGAGACTTTACCACTGTCAA	2333
Db	1959	GGCTAACACACTCATCTTCCATCTCTCTGCACTCGAGGAGGCGCTGCATTCACCGCTGCACA	2018
Qy	2334	GATTGTCAACTACACATGTACGCTTCGGGCCCTCGGGGCGCTCAGCAATGCTTCCCTGGTCCAC	2393
Db	2019	AATCATTAATCAACCATCGCTGTAATCTGTGTGGCTCGCAGAAATGCTTCTTTGGTTCT	2078
Qy	2394	CTACTATACAGCTCTGGAGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTA	2453
Db	2079	GTACCTTCAAAACATCAAGAACTCACATGTTGGTTCAAGGCTTAGAAACCAACACCAATA	2138
Qy	2454	CGAGTTGGCGGTACAGTCCCAACGAGTGGATATGGATGGGCCCTTTGGCTCCGTCGTAGA	2513
Db	2139	CGAATTTGCGGTTCGAATTACATGTGATCAGCTTTCCAGTCTTTGGAGCCCTGTAGTCTTA	2198
Qy	2514	ACGCTCCACCTCGCTGACCGGCTTCAACACACTCTTTCTGACCTCGGCTGAGCCCCCT	2573
Db	2199	CCATTCTACTCTTCCAGAGACACGACGAGGCCACCAAGTTGGAGTAAAGTGACATTAAT	2258
Qy	2574	GACACCATCACCGTTGGTTTACACTGGTGTCCCCCAACGAGGCCCAATGGTGAGATTGT	2633
Db	2259	AGAGGATGACACTGCGCCCTGGTTTCTTGGAAACCCCTGTATGGCCGAGAAACAGTTGTGAC	2318
Qy	2634	GGAGTATCTAATTTCTCTACAGCAACAAACACACCCAGCCCGGAACACACAGTGGACACTGCT	2693
Db	2319	CCGCTATACTATCTTATATGCACTAGGAAGGCTCGAATTTGACGAGAGTGGCAGGTCCT	2378
Qy	2694	CACCAAGAGGGGAAAACATCTTCAGTGAGAGGTCCATGGCCTAGAGAGTGACACTCGGTA	2753
Db	2379	ACACCGTGAAGGGGCAATAACCATGGCTTTGCTAGAAACATTTGGTAGCAGGNAATGTGTA	2438
Qy	2754	TTTCTTCAAGATGGGAGCCCGCACAGAGGTGGGGCCCTGGGCCCTTTTCCCGCTTGACGGA	2813
Db	2439	CATTGTCAAGATATCTGCATCCAAATGAGGTGGGAGGAGGACCCCTTTTCAAAATCTGTGGA	2498
Qy	2814	TGTGATTACTCTGCAAGAGACATTTCCAGACTC	2846
Db	2499	GCTGGCAGTACTTCCAAAGGAAACCTCTGAATC	2531

RESULT 12
ABX71447

ABX/144/
ID ABX71447 standard; cDNA: 3874 BP.

AC ABX71447;

DT 06-MAR-2003 (first entry)

DE Human cDNA encoding a novel tyrosine phosphatase.

Human; ss; gene; chromosome 15; chromosome 3; tyrosine

to virus family protein; gene sheaf; NHR; novel human protein.
XX

OS Homo sapiens.

PN US6465632-B1.

PD 15-OCT-2002.

08-JUN-2001;

AA
PR 09-JUN-2000; 2000US-210607P.AA (LEXI-) LEXICON GENETICS INC
PA (LEXI-) LEXICON GENETICS INC

XX Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
XX WPI; 2003-147071/14.
PT Novel isolated nucleic acid which encodes a novel human protein that
PT shares sequence similarity with animal phosphatases, that is useful for
PT generating antibodies, and as reagents in diagnostic assays -
XX Disclosure; Column 107-112; 58pp; English.
XX The invention relates to an isolated nucleic acid molecule comprising a
CC nucleotide sequence encoding a novel human protein (NHP) that shares
CC sequence similarity with animal phosphatases (in particular tyrosine
CC phosphatases and is a member of the Ig superfamily). The disclosed NHP
CC polynucleotide sequences (from genes located on either chromosome 15
CC or chromosome 3) are useful for identifying coding sequences and in the
CC identification of biologically relevant splice junctions. The NHP
CC polynucleotide sequences are useful in gene therapy, for detecting of
CC mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NHP.
CC Complementary sequences of the NHP polynucleotides used in conjunction
CC with PCR to screen libraries, isolate clones and prepare cloning and
CC sequencing templates. Labeled NHP nucleotide probes can be used to screen
CC a human genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests. The probe sequences also have use in defining and
CC monitoring both drug action and toxicity. Oligonucleotides
CC complementary to NHPs may encode or act as NHP antisense molecules, or
CC may be used as part of ribozyme and/or triple helix sequences.
CC The present sequence encodes an NHP.
XX
SQ Sequence 3874 BP; 1077 A; 938 C; 897 G; 960 T; 2 other;
Query Match 5.1%; Score 316.6; DB 25; Length 3874;
Best Local Similarity 47.4%; Pred. No. 3.6e-71;
Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6;
QY 300 AGACAGGATGATGAGAAAGCTCTTAGGATCTGGAGGTGCACTGAGGGGAGCTATTCCTG 359
DB ATACATCAGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 528
QY 360 TCTGGCCACAGCCGCTAGGAGTGTGGCCAGCCAGGTTGCTGTTCAAGTTCAGCTCCAC 419
DB 529 CTTGGCAATGAAACAATATGGAGCCATCTTAGTCAAAAAGCTCATCTTGCCTTATCAAC 588
QY 420 ACTCGAAGACTTCTCTGACCCCGAGTCCAGATTTGTGGAGGAGAAACGGGACAGCACG 479
DB 589 TATTTCTGCATTTGMAAGTCCAGCCAAATTTCCACTGAGGTCACGAAAGGTGAGTTGCTG 648
QY 480 CTTTGAATGCCACACCAAGGGCTTTCAGGCCCATCATTTACTTTGGGAAAGGACAGGT 539
DB 649 ATTTGATGCAAGATTTTCATCCCAACCTCTGCGAGTCATTAACATGGAGGTTTCAATCGGAC 708
QY 540 GACGCTGCTGAG---GAGCCCGGCTCATCATCTTCCCAAGTGGCTCTCCAGATCCT 596
DB 709 AACTCTACCTATGACTATGACAGGATACTGCCCTACCAACAGGAGTATTGACAGATCTA 768
QY 597 AGATGTCCAGGAGTATGACAGGCTCTTACCGCTGGTGGCCCAACCAATTCAGCCCGCCA 656
DB 769 TGATGTCCAGCNAAGGATTTCTGMAATATCTGTTGATTTGCTGCCACTGTAGCCCAACG 828
QY 657 ACGATTCCAGCAGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTTGAGGCTACACAG 716
DB 829 ACGTAAAGTATGAGGAGGCTCGCTCACTGTGATTTCCAG-----CTAAGGAGTCAAAATC 882
QY 717 GGGGACAGATGTGCTATTTGGGAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 776
DB 883 CTTCCACACACCAACCAATATAGCAGGTCCACGAAACATTAACCAATCTCTTCATCAGAC 942
QY 777 TGTAGTGTGAGGTGCGGCTCTGCTGACCCCAACCCCTTTTGTGTCTCGGCTCCGACA 836

DB 943 TGTAGTTTTGGAAATGCATGGCCACAGGAAATCCCAAAACCAATCATTTCTTGGAGCGCCT 1002
QY 837 GGATCGAAAGCCTAT---CTCCACGGATGTCTGCTGCGCGGACCAATCTACTCAT 893
DB 1003 TGTATCAAAATCCATTTGATGTCTTTAATCTCGGGTACTTTGAAATGGTAATCTCATGAT 1062
QY 894 CGCCAGCGCGCAGCCTCTGGGCACTCTGGAGTCTATGTCTGCCGAGCAACCAAGCCCTCAC 953
DB 1063 ATCTGATGTGAGGCTACAAACATGCTGGAGTATATGTTGTGCGGCGCACTACCCCTGGCAC 1122
QY 954 GGTGATCTTGGCACTGGCGGCTGAGCTCGAGTCTCGAGTCTGTGTCGCCCGCAGCATCTCGCA 1013
DB 1123 ACGCAACTTTACAGTTGCTATGCGCAACTTAACTGATGATTTAGTCTCTCTTCAATTTGTTGA 1182
QY 1014 GGCACCCGAGGCGCTCTCGGAGCGCGGCGCAGCAGCGCGCTTGTGTGCGGCGGCTC 1073
DB 1183 ATGGCCAGAAAGTTTAAACAGGCTCTGAGCTGGCACTGCTCGATTTGTGTGTCAGGAGA 1242
QY 1074 CGGGGAGCCACCGGCCCGCTGCACTGGCTGCACGACGCGGATCCCGTTTGCACCCAAATGG 1133
DB 1243 AGGAATCCCTCTCCAAAGATGTCTGTTGTTGAAAATGGAAGAGATACATTCGAATGG 1302
QY 1134 GCGGTCAAGGTGCAAGGGGGTGGCGGAGCTTGGTCTCATCTCAGATCGGCTCGCAGGA 1193
DB 1303 TAGAATTTAAATGTACACAGT-----AAATTTGGTAAATTAACCAAGATTTATTCCTGAAGA 1356
QY 1194 CGCTGGCTACTACCACTGGTAGCAAGAAACAGCGCGGAACTGCTGTGCGCTGCGCC 1253
DB 1357 TGATGCTATTTATCAGTGTGATGGCTGAGAAATAGCAAGGATCTATTTATCTAGAGCCAG 1416
QY 1254 CTTGCGGTAGTGTGTCGCGAGGGGCTGCCAGCGCCCGCACTCGGGTCAACAGCCACGCGC 1313
DB 1417 ACTGACTGTGATGTGTGAGAGACAGACCCAGTGTCTCCCTATATGTATGATGCTGGAAC 1476
QY 1314 GCTGAGAGCTCTCTGTGCTGGTGGCTGGGAGCGGCTGAGTGTGCAAGGAGGAGCAAT 1373
DB 1477 CATGTCAAGCTCAGCCATCTTTTATAGCTTGGGAGAGGCCACTTTATTAATTCAGACAAAGT 1536
QY 1374 CATTCGCTCTCTCTTCACTACCAAAAGCAAGGGGAGTGGCAATGTGGAGTACCAGTT 1433
DB 1537 CATTCGCTATTTCTGTACACTACATGAAGAGAGAGGTTTAAATATGAAGAGTATCAGGT 1596
QY 1434 TGCAGTAAACAAATGACACCAAGAGCTGCGAGGTTTGGGACCTTGGGACCCCAACCAACGATTA 1493
DB 1597 AGTCATCGGAATGACAACTCATTTATTTATGATGATTTAGAGCTTCCGCGGCAATTA 1656
QY 1494 TGAGTTTACGTGGTGGCTTCTCCAGCTGGGGGCGAGCGAACTCCAGGCCAGCCCT 1553
DB 1657 TACTTTCTACATTTGTAGCATATATGCCAATGGGAGCCAGAGATGCTGAGCCATGTGAC 1716
QY 1554 GGTGATACATCGGAGATGTCGCCAGCGCAGCAGCCCGAGCTTACCTTGTCCAGCCCA 1613
DB 1717 ACAGAAATCTCTAGAGGATGTTCCCTCTGAGACCTCTCTGAAATTTAGTTTGAAGTCTGAG 1776
QY 1614 CCGCTCGACATCAGGGTGGCATGCTGCCCTTCCCTGCGCTCCAGCTCAGCAATGGAAGT 1673
DB 1777 TCCCACTGATATCTCATCTCTGCTGCCAATCCAGCGCAATATCGCGGGGCGCAAGT 1836
QY 1674 GCTGAAGTACAAGATAGAGTACGTTTGGGGAAGGAAGATCAGGTTTTTCTCCACCGAGGT 1733
DB 1837 GGTGCTGATCGCTTGTCTTTCCGCTAAAGTACTGAGAAATTCAAATCCAAGTTCTCGAGCT 1896
QY 1734 GCTTGAATGAGACAACTTACGTTAACTCATCTTCAGCCCAACCAAGTGTACCGAGT 1793
DB 1897 CCGGGGAGCCACGCAATGAGTACCTTTTGAAGGCGCTGAAACCTGACAGTGTCTACCTGGT 1956
QY 1794 CCGGATTTTCAGCTGGCAGCTGGCGCTGGCTATGAGTCCCTTCTCAGTGTGAGCAGCAG 1853
DB 1957 TCGATTTACTGTGCCACAGAGTGGGGCTGGGAGAGTCAATCAGTATGAGCTTCATCAGTAG 2016
QY 1854 GACACCTGGTGTGCAACCAACAGAGCCATGTTCCCTTTGCGCTTCGAGAAATGGAAGTGTAG 1913

Db 2017 GACGCCAAAGCTACAGCGGTGAAGCCCTTAAGTCTCCAGATTGTCATTTGGAGCCTCT 2076
Qy 1914 GGCAGAGATGAGTCCCTGGTGTGTCATGGCAGCGCCCTCACCCACCCAGATCTC 1973
Db 2077 GAACCTGTACCACTTCTGTGAGGTGGCAGCAAGATGATAGAGACACAGCTGTATTCA 2136
Qy 1974 TGGATACAACTCTACTGGGAGAGGTGGGAACAGAGAGAGGAGGATGTGACCGCCC 2033
Db 2137 GGGCTACAAAGCTGTACTACAAGGAAGAGGGCAGCAGAGAA----- 2178
Qy 2034 CCCAGGGGCTGTGGAGATCAAGCTTGGGACGTGGGCGCTGGCGCTGAAGAGAAAGT 2093
Db 2179 -----TGGGCCCAATTTCTTGGATACCAAGGA 2205
Qy 2094 GAAGCAGTATGAAGTCAAGTTCCTGGCAGGCGGTACGAGGTGAAGCTCTGTAGC 2153
Db 2206 CCACTCTATCTCTCACTGGCTTAGACCCCAAGAAATATCATGTGAGACTCTCTGGC 2265
Qy 2154 TTTCAACAAACACGAGGACGGCTAGCTGTGTGTGGGAAGGGCAGAGAGAGCGCC 2213
Db 2266 TTAACAACAATAGACGATGGCTATCAGGCGAGATCAGACTGTGAGCAC---TCCAGGATG 2322
Qy 2214 CACGCCAGCTGCTATCCAGAGGGGGCCACCGCTGCTCTGCCCATGTCCAGCGAGA 2273
Db 2323 CGTGTCTGTGATCGATGGTCCCTCCACCAACCCACCCACCATCTCTATGCGAA 2382
Qy 2274 GTCAACAGCTCCACTTCATTTGGCTTTCGGTGGGAAGACAGACTTTTACCACTGTCAA 2333
Db 2383 GGCTAACACCTCATCTTCATCTCTGCTGCTGCTGAGGAGGCTGATTCACCGCTGCACA 2442
Qy 2334 GATTGTCAACTACACTGTACCTTGGCCCTTGGGGCTCAGGAATGCTTCCCTGGTCA 2393
Db 2443 AATCAATTAACACCACTCCGTGTAATCTCTGTTGGCTGCGAATGCTCTTTGGTTCT 2502
Qy 2394 CTACTATACAGCTCTGAGAGAGATCTCATTTGGGGCTGGAACCAATTTACCAAGTA 2453
Db 2503 GTACCTTCAACATCAGAAATCTCAATGTTGGTTCAAGGCTAGAACCAACCAACCAATA 2562
Qy 2454 CGAGTTTGGGTACAGTCCACGAGGTGGATATGGATGGGCGCTTTGGCTCGTGTAGA 2513
Db 2563 CGAATTTGGCTTGAATATCATGTGGATCAGCTTTCAGTCTTGGAGCCCTGTAGTCTA 2622
Qy 2514 ACGTTCACCTGCTGACCGGCTTCAACACCTCTTCTGACCTGCGCTGAGCGCCCT 2573
Db 2623 CCATTTCTACTCTTCAGAGACACGAGCGGCGCCACCACTTGGAGTAAAGTGACATTAAT 2682
Qy 2574 GACACCATCCAGCTTGGTGTACTGTGTGTCCTCCCAAGGAGCCCAATGTGTGATGT 2633
Db 2683 AGAGGATGACACTGCTGCTGTTCTTGGAAACCCCTGATGGCCCGAGAAACAGTTGTGAC 2742
Qy 2634 GGAGTATCTAATTTCTACAGCAACACACACCGCCCGAACCAGTGGACACTGCT 2693
Db 2743 CCGCTATCTATCTTATATGATCTAGAAAGGCTTGGATGAGAGAGTGGCGGCTT 2802
Qy 2694 CACCACAGAGGGAACATCTTTCAGTGCAGAGGTCCATGGCTAGAGAGTGAACCTCGGTA 2753
Db 2803 ACACCGTGAAGGGCAATAACATGGCTTCTAGAAACCTTGTGACGAGAAATGTGTA 2862
Qy 2754 TTTCTTCAAGTGGAGCCCGACAGAGTGGGCGCTTGGGCTTTTCCCGCTTGCAGGA 2813
Db 2863 CATTTCAAGATATCTGATCCATGAGGTGGGAGAGGACCTTTTCAAAATCTGTGGA 2922
Qy 2814 TGTGATTAATCTGCAAGAGACATTTCTAGATC 2846
Db 2923 GCTGGCAGTACTTCCAAAGGAACCTCTGNAATC 2955

RESULT 13

RAD36777 standard; cDNA; 2700 BP.

XX AAD36777;

AC AAD36777;

XX

DT 30-OCT-2002 (first entry)
XX Human immunoglobulin superfamily protein (IGSFP)-1 cDNA.
XX Human; immunoglobulin superfamily protein-1; IGSFP-1; asthma;
XX immune system disorder; acquired immune deficiency syndrome; AIDS;
XX atherosclerosis; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; developmental disorder; renal tubular acidosis;
XX anemia; muscle disorder; cardiomyopathy; myocarditis; cancer;
XX cell proliferative disorder; arteriosclerosis; hepatitis; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..2382
XX /tag= a
XX /product= "IGSFP-1"
XX sig_peptide 1..105
XX /tag= b
XX mat_peptide 106..2379
XX /tag= c
XX /product= "Mature IGSFP-1"
XX MO200240671-A2.
XX 23-MAY-2002.
XX 15-NOV-2001; 2001WO-US44974.
XX 16-NOV-2000; 2000US-249645P.
XX (INCY-) INCYTE GENOMICS INC.
XX Baughn MR, Lu DAM, Yue H, Elliott VS, Thangavelu K, Ramkumar J;
XX Lu Y, Lo TP, Gururajan R, Gandhi AR, Arvizu C, Yao MG,
XX WPI; 2002-519384/55.
XX P-PSDB; AAE14781.
XX Novel human immunoglobulin superfamily polypeptide, useful in
XX diagnosis, prevention or treatment of immune system, neurological,
XX developmental, muscle and cell proliferative disorders -
XX Claim 5; Page 116-117; 122pp; English.
XX The present sequence is human immunoglobulin superfamily protein
XX (IGSFP)-1 cDNA. The IGSFP polypeptide and polynucleotide are useful
XX for diagnosing, treating or preventing disorders associated with aberrant
XX expression of IGSFP e.g. immune system disorders (e.g. acquired immune
XX deficiency syndrome (AIDS), asthma, atherosclerosis, psoriasis, uveitis),
XX neurological disorders (e.g. Alzheimer's disease, Huntington's disease,
XX dementia, Parkinson's disease), developmental disorders (e.g. renal
XX tubular acidosis, epilepsy, anaemia), muscle disorders (e.g.
XX cardiomyopathy, myocarditis), or cell proliferative disorders (e.g.
XX arteriosclerosis, cirrhosis, hepatitis, and cancer). The polypeptide and
XX polynucleotide are also useful for assessing the effects of exogenous
XX compounds on their expression. The polypeptide is useful in drug
XX screening techniques, to analyse the proteome of a tissue or cell type,
XX as elements on a microarray. The polynucleotide is useful for creating
XX knock-in humanised animals or transgenic animals to model human diseases,
XX in somatic or germline gene therapy, to generate a transcript image of a
XX tissue or cell type, for detecting differences in the chromosomal
XX location due to translocation, inversion among normal, carrier or
XX affected individuals, and as hybridisation probes for mapping naturally
XX occurring genomic sequences.
XX Sequence 2700 BP; 528 A; 859 C; 815 G; 498 T; 0 other;

Query Match 5.1%; Score 316.4; DB 24; Length 2700;

Best Local Similarity 55.0%; Pred. No. 3.3e-71;

Matches 692; Conservative 0; Mismatches 551; Indels 15; Gaps 3;

Qy 342 TGAGGGCAGCTATTCTCTGCTGCGCCACAGCCCGCTAGGATGTGTGCGCCAGCGTTGC 401

Db 333 TGAAGGTGACTATGAGTGTGTGGCCAGAAACCGCTTTTGGGCTGTGTGGTTCAGCCGGAAGGC 392
QY 402 TGTGGTCAAGCTTGCCACACTCGAAGACTTCTCTGTCACCCCGAGTCCAGATGTGGA 461
Db 393 TCGCATCCAGCTGCAACCAATGTCGACCTTCACAGTGCATCCCGAGGCCACCGTGGGTGA 452
QY 462 GGAGAAACGGGACAGCAGCTTTGAAATGCCACACCAAGAGGCTTTCAGAGCCCCCATCATNTAC 521
Db 453 GGAAGGTGTGTGGCCCGCTTCAGTGCACAAATCCATGGCTTCCCAAACCCCTGATCAC 512
QY 522 TTGGGAAAGGACCAAGT---GACCGTGTGAGGAGCCCGGCTCATCACTTTCCTCAA 578
Db 513 TTGGGAGAGAACAGAGTCCCAATTGACACGGCAATGAGAGGTACACATTTGTCGCCAA 572
QY 579 GTGGCTCTCCACATCTAGATGTCCAGGACAGTGTATGACGGCTCTCAACGCTCGTGTGC 638
Db 573 GGGGCTCTGCAGATCAGAGCTTCAGCTGAGGACGTGGCATTTTCACTGTGTGGC 632
QY 639 CACCAATTCAGCCCGCCAAACGATTCAGCCAGGAGGCTCTGCTCACTGTGTGGCTTCAGAG 698
Db 633 CTCAACATCGCCAGTATCCGGATCAGCCAGCGGGCCAGGCTCACTGTGTGAGGCTCGGG 692
QY 699 GTCTTTGGAGGCTACAGGGGAGGATGTGTTCATTTGTGGCAGGCCCCAGAGAACACAC 758
Db 693 CTCCTGGGGCTACA-----AGGAGCCAGCCATCTCTGTGGGGCTTGAGAACCTCAC 743
QY 759 GGTAGTGTCTGGACAGAAATGATGATGAGTGTGGTGTGCTGTGACCCCAACCCCTTT 818
Db 744 CTGACAGTGCACACCGCGTGTGAGTGTGTGCGCACGGGACCCGGGCCCAT 803
QY 819 TGTGCTCTGGGTCCGACAGAGTGAAGCTTATCTCCACGGATGTCTATC---GTTCTGGG 875
Db 804 TGTGCTCTGGAGCGCTGATGTTGCTGCTATCGGGGTGGAGGCGATCCAGTGTCTGGG 863
QY 876 CCGGACCAATCTACTCATGCCAGCGCGCAGCCCTCGGCACTCTGAGTGTCTGTCTGCGG 935
Db 864 CACAGGAACCTCATCTCTGACCGTGCAGCTGCGGTCCAGCACCTCTGGGCTTACGTCTGTGC 923
QY 936 AGCCAAACAGCCCTCAGCGGTGACCTTCGCCACTGCGGCTGTGAGCTCGAGTGTCTGC 995
Db 924 AGCCAAACAGCTGGCACCGGGTGAGAGAACGGCACAGGGCCGCTGTGTGTGACG 983
QY 996 TGCCCCAGGACATCTCGCAGGACCCCGAGCGCTCTCGCGACCGCGGCGCAGCACCGCGG 1055
Db 984 CCAGCTGAGTTGTGACGATCCCGAGTCCATCTCAGGCCAGCTGGGACACAGCCAT 1043
QY 1056 CTTGCTGTGCGGGCTCGGCGAGCCAGCGCCGCGCTGCACTGGCTGCAAGAGGAT 1115
Db 1044 GTTCACCTGCAAGCCAGGGTGAGCCAGCGCTCATGTCACTGGCTGAAAATGGACA 1103
QY 1116 CCGTGTGCGACCAATGGGCGGTCAAGTGCAGGGCGGTGGCGGAGCTTGTCTATCAC 1175
Db 1104 GTTGTGCGGCGCAGAGGCGCACGTCAAGGCTCAAGATATAACAGACACACTGACCATTT 1163
QY 1176 TCAGATCGGCTCGCAGAGCGCTGGCTACTACAGTGCCTAGCAGAAAACAGCGCGGGAAC 1235
Db 1164 TGAATCGGCTCTGAGAGTGAAGCCATTTATAGTGTGTGGCGAGAACAGTGGGGCTC 1223
QY 1236 TGCCTGTGCGGCTCGCCCTGCGGTAGTGTGCGCGAGGGCTGCCAGAGCGCCCGAC 1295
Db 1224 ATCACAGGCGAGTGCACCGGTGACCGTACTGTGGGCTGAGGGGCTCCCGGGCTCCCG 1283
QY 1296 TCGGFTCACAGCCAGCGCTGAGCAGCTCTCTGTGCTGTGGTGGCTTGGAGGGGCTGA 1355
Db 1284 CAATGTGCGGGCAGTCTCTGTGTCTTCCATGAGGTGCGGTGTCTCTGAGAGGCGCT 1343
QY 1356 GTTGCACAGGAGCAATCATTTGGCTTCTCTTCACTACCAAAAGGAGGAGGTGGA 1415
Db 1344 GGCCAAACAGAGGATCATCGGCTACGTCTGCAATCAGGAAGGCTGTGACCCACC 1403
QY 1416 CAATGTGAGTACCACTTTGCAAGTAAACAAATGACACACAGAGCTGCAGGTTGGGACCT 1475

Db 1404 GGAGCTGGAGTATCAGGAGGCGAGTCAGCAAGAGCACCTTTGAGCACTGTGTGAGCACT 1463
QY 1476 GGAACCCCAACACGGAATTATGAGTTCTAGCTGTGTGGCTTACTCCCAAGCTTGGGGCCAGCG 1535
Db 1464 GGAGCCCTCCACAGCGCTACAGTTTCTACATCAAGGCTTACACACCAAGGGGGCCAGCTC 1523
QY 1536 AACCTTCAGCCAGCCCTGGTGTGCATACACTGACATGTCCCAAGCGAGCAACCCAG 1593
Db 1524 AGCCTCTGTGGCCACCTTAGTCAGCACCTGTGGTGAAGCCCTGTGCCACCCCACTG 1581

RESULT 14

ABX71435

ID ABX71435 standard; cDNA; 3453 BP.

XX AC ABX71435;

XX DT 06-MAR-2003 (first entry)

XX DE Human cDNA encoding a novel tyrosine phosphatase, NHP4.

XX KW Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;

XX KW Ig super family protein; gene therapy; NHP; novel human protein.

XX OS Homo sapiens.

XX PN US6465632-B1.

XX PD 15-OCT-2002.

XX PF 08-JUN-2001; 2001US-0877730.

XX PR 09-JUN-2000; 2000US-210607P.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;

XX PI Zambrowicz B, Sands AT;

XX XX WPI; 2003-147071/14.

XX DR P-PSDB; ABU54203.

XX PT Novel isolated nucleic acid which encodes a novel human protein that shares sequence similarity with animal phosphatases, that is useful for generating antibodies, and as reagents in diagnostic assays

XX PS Claim 1; Column 31-36; 58pp; English.

XX CC The invention relates to an isolated nucleic acid molecule comprising a nucleotide sequence encoding a novel human protein (NHP) that shares sequence similarity with animal phosphatases (in particular tyrosine phosphatases and is a member of the Ig superfamily). The disclosed NHP polynucleotide sequences (from genes located on either chromosome 15 or chromosome 3) are useful for identifying coding sequences and in the identification of biologically relevant splice junctions. The NHP polynucleotide sequences are useful in gene therapy, for the diagnosis of mutant NHPs or inappropriately expressed NHPs, for the treatment of disease, for screening drugs effective in treatment of symptomatic or phenotypic manifestations of perturbing the normal function of NHP. Complementary sequences of the NHP polynucleotides used in conjunction with PCR to screen libraries, isolate clones and prepare cloning and sequencing templates. Labeled NHP nucleotide probes can be used to screen a human genomic library which is helpful for identifying polymorphisms, determining the genomic structure of a given locus/allele and designing diagnostic tests. The probe sequences also have use in defining and monitoring both drug action and toxicity. Oligonucleotides complementary to NHPs may encode or act as NHP antisense molecules, or may be used as part of ribozyme and/or triple helix sequences.

XX CC The present sequence encodes an NHP.

SQ Sequence 3453 BP; 979 A; 821 C; 769 G; 883 T; 1 other;

Query Match

5.1%; Score 315.4; DB 25; Length 3453;

Best Local Similarity 47.4%; Pred. No. 6.9e-71; Matches 1210; Conservative 0; Mismatches 1271; Indels 72; Gaps 6;			
QY	300	AGACAGCATGATGAGGAGCTTTAGGATCTGGAGGTCTACTGAGGCGAGCTATTCCTG	359
DB	279	ATAATCATGATGAGTGGAGGCGGAGGAGAGCTCCGATGAGGATTTATTCAGTG	338
QY	360	TCTGGCCACAGCCCGCTAGGAGTGGTGGCCAGCGGTGTGTGTGCTCAAGCTTGGCAC	419
DB	339	CTTGGCAATGAACAATAATGAGGCCATTTCTTAGTCAAAAAGCTCATCTTGCCTTATCAAC	398
QY	420	ACTCGAAGACTTCTCTGACCCCGAGTCCAGATTTGTGGAGGAAGGACAGCAGC	479
DB	399	TATTTCTGCAATTTGAAGTCCAGCCAAATTTCCACTGAGGTCCACGAAGGTGGAGTTGCTCG	458
QY	480	CTTTGAATGCCACACCAAGGGCTTCCAGCCGCCCATCTACTTGGGAAAAGGACCAAGT	539
DB	459	ATTTCGATGCAAGATTTATCCCACTCTGCGAGTCAATACATGGGAGTTCAATCGGAC	518
QY	540	GACCGTGGCTGAG---GAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCCT	596
DB	519	AACTCTACTATGACTATGAGCAGGATAACTGCGCTACCAACAGGAGTATTGCGATCTA	578
QY	597	AGATGTCAGGACATGATGAGGCTCTCTACCGCTGCGTGGCCACCAATTCAGCCGCCA	656
DB	579	TGATGTCAGCCAAAGGGAATCTGGAAATTTATCGTTGTATTTGCTGCCACTGTAGCCACCG	638
QY	657	ACGATTCAGCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAG	716
DB	639	ACGTAAAGTATGGAGGCTCGCTAACTGTGATTCAG-----CTAAGGAGTCAAAATC	692
QY	717	GGGGCAGGATGTGTGCTATTTGGCAGCCCGCAGAGAACACACGCTAGTGTCTGGACAGAA	776
DB	693	CTTCCACACACCAACAAATTATAGCAGGTCACAGAAACATACAACTCTTTCATCAGAC	752
QY	777	TGTAGTATGAGTGGTGGCTCTGTGACCCCAACCCCTTTTGTGTCTGGTCCGACA	836
DB	753	TGTAGTTTGGAAATGCAATGGCCACAGAGAAATCCCAACCAATCAATTTCTTGGAGCGCT	812
QY	837	GGATGGAAGCTAT---CTCCACGGATGTCTATCGTTCTGGCCGGACCAATCTACTCAT	893
DB	813	TGATCACAAAATCCATTTGATGTCTTTAACTCTGGGTACTTGGAAATGTGAATCTCATGAT	872
QY	894	CGCCAGCGCAGCGCTCGGCACTCTGGAGTCTATGTCTGCCAGCAACCAAGCCCTCAC	953
DB	873	ATCTGATGTCAGGCTAGACATGCTGGAGTATATGTTTGTGGGCCACTACCCCTGGCAC	932
QY	954	GGTGACTTCGCACTGGCGCTGTGAGCTCGAGTGTGCTGCCCGCCAGCCATCTCGCA	1013
DB	933	ACGCAACTTTACAGTTGTATGGCAACTTTAACTGTATTTAGCTCTCTTCAATTTGTGA	992
QY	1014	GGCACCAGGCGCTCTGGCGACCGGGCCAGCACCGCGCTTCTGTGTGGCGGCGTC	1073
DB	993	ATGGCCAGAAAGTTTAAACAGGCTCTGAGCTGGCACTGCTCGAATTTGTGTGAGGCGACA	1052
QY	1074	CGGGGAGCACCGCCCGCTGCACTGTGCTGCACACCGGATCCCGTTGGACCCAAATGG	1133
DB	1053	AGGAATCCCTCTCCAAAGATGTCAATGGTTGAAATGGAAGGAAGATACATTCGAATGG	1112
QY	1134	CGCGCTCAAGGTGACAGGCGGTGGCGAGCTTGTGCTATCACTCAGATCGGCTGCGAGA	1193
DB	1113	TAGAATTAATGTACACAGT-----AAATTGGTAATTAACACAGATTTATTCGTGAAGA	1166
QY	1194	CGCTGGCTATACAGTGGTGAAGAAACAGCGCGGGAACTGCTGTGCGCTGCGCC	1253
DB	1167	TGATGCTATTTATCAGTGTGATGGCTGAGAAATAGCCAAAGGATCTATTTTATCTAGAGCCAG	1226
QY	1254	CTGCGGTGATGTGGCGAGGGCTGCCAGCGCCCGACCTCGGCTCAGAGCCACGCC	1313
DB	1227	ACTGACTGTAGTGTGATGACAGAGACAGACCCAGTCTCCCTATATATGATGACTGTGAAC	1286
QY	1314	GCTGAGCAGCTCTCTGTGCTGGTGGCTGGGAGCGGCTGTAGTTGCAAGGAGCAAAAT	1373

DB	1287	CATGTCAGGCTCAGCCATTTCTTTAGCCTGGGAGAGGCCACTTTATTAATTCAGACAAAGT	1346
QY	1374	CATTGGCTTCTCTCTTCACTACCAAAAGGCAAGGGAGTGACCAATGTGGAGTACCAGTT	1433
DB	1347	CATTGGCTTCTCTCTGTA CACTACATGAAGCAGAGGTTTAAATTAATGAAGAGTATCAAGT	1406
QY	1434	TGCAGTAAACAAATGACACACAGAGCTGCGAGTTGCGGACCTGGAACCCAAACAGGATTA	1493
DB	1407	AGTCATCGGAAATGACACAACTCATTTATTTATTTGATGACTTAGAGCCTGCGCAGCAATTA	1466
QY	1494	TGAGTTCTAGTGTGGCTTACTCCAGCTGGGGGCGAGCGAACCTCCAGCCCGCCT	1553
DB	1467	TACTTTCTTACATTTGTAGCATATATGCCAATGGGAGCGCAGCATGTCTGACCATGTGAC	1526
QY	1554	GGTGATACACTGACGATGTCCCGCAGCGCAGCACCCCGAGCTTACCTTGTTCAGCCCCAA	1613
DB	1527	ACAGAACTACTAGAGGATGTCCCTCGAGACTCTCTGAAATTAATTTGACAACTGCGAAG	1586
QY	1614	CCCTCGGACATCAGAGGTGGCATGTGCTGCCCTCGCTCCAGCTCGAGCAATGACAGGT	1673
DB	1587	TCCCACTGATATCTCATCTCTGGCTGCCAATCCAGCCAAATATCGCGGGGCCAAGT	1646
QY	1674	GCTGAAGTACAAAGATAGGTACGGTTTGGGGAAGGAGATCAGGTTTCTTCTCAAGAGGT	1733
DB	1647	GGTGTGTATCGCTTGTCTTTCGCTTAAGTACTGAGAAITTCAAATCCAAAGTTCTGGAGCT	1706
QY	1734	GCCTGGAAATGAGACACAACTTACGTTAAACTCTCACTTCAGCCAAACAAAGTTGACCGAGT	1793
DB	1707	CCCGGGAGCAGCATGATGATGATCTTTTGGAAAGGCTGAACTGACAGTGTCTACTCTGCT	1766
QY	1794	CCGGAATTCAGCTGGCACTGGCGCTATGGAGTCCCTTCTCAGTGGATGACGACACAG	1853
DB	1767	TGCGATTACTTGCTCCACAGAGTGGGCTGGGAGAGTCACTAGTATGAGCTTCACATAG	1826
QY	1854	GACACTGTGTGACAAACACAGAGCATGTTCCCTTGTGCCCTCGCAGAAATTTGAAGGTGAG	1913
DB	1827	GACGCCCAAGCTACAAGCGTGAAGCCCTTAAGTCTCAGAGTTGCAATTTGGAGCCTCT	1886
QY	1914	GGCAAGATGAGTCCCTGTGTGTCTATGGCAGCGCCCTCTCACCCACCCACAGATCTCTC	1973
DB	1887	GAACTGTACCACTATTTCTGTGAGTGGCAGCAAGATGTAGAGGACACAGCTGTCTATTCA	1946
QY	1974	TGGATACAAACTCTACTGGGAGAGGTGGGAAACAGAGAGGAGGAGGAGTGTGACCGGCC	2033
DB	1947	GGGCTACAAGCTGTACTACAAGGAAGAGGCGCAGCAGGAGAA-----1988	
QY	2034	CCCAGGGGCTCGTGGAGATCAAGCTTGGGAAGTGGGCGCTGGCGCTGGAAGAGAAAGT	2093
DB	1989	-----TGGGCCCATTTTCTTGGATACCAAGGA 2015	
QY	2094	GAAGCAGTATGAACCTGACCCAGTTAGTCCCTGGCAGGCGGTACGAGGTGAAGCTCGTAGC	2153
DB	2016	CCTACTCTATCTCTCAGTGGCTTAGACCCCGAAGAAATAATCATGTGAGACTCCTGGC	2075
QY	2154	TTTCAACAAACACGAGGACGGCTACGCTGTGTGTGGAAAGGCGCAAGACGGAAGCGGCC	2213
DB	2076	TTACAAACACATAGACGTGCTATCAGGACAGATCAGACTCTCAGCAC---TCCAGGATG	2132
QY	2214	CACGCGACAGCTGCTATTCAGAGGGGGCCACCGCTGCTCTTGTGCCCATGTCCAGCGAGA	2273
DB	2133	CGTGTCTGTTGCTGATGCGATGTCTCTCTCCACCAACCCACCATCTCTATGCGAA	2192
QY	2274	GTCAACAGCTCCACTTCCATTTGGCTTGGGTGGAGAGGAGGAGGAGGAGGAGGAGGAGG	2333
DB	2193	GGCTAAACACCTCATCTTCCATCTTCTCTGCACTGGAAGGAGGCTTGCATTCACCGCTGACA	2252
QY	2334	GATTGTCAACTACACTGTACGCTTTCGGCCCTTGGGGGCTCAGGAATGCTTCTCGGTAC	2393
DB	2253	AATCATTAACATACACATCCGCTGTATCTTGTGGCTGAGAAATGCTTCTTGTGTTCT	2312
QY	2394	CTACTATACAGCTCTGGAGAGGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTA	2453
DB	2313	GTACCTTCAACATCAGAAACTCATATGTTGTTCAAGTCTTAGAACCCAAACACCAANAATA	2372


```
Db 922 AGGCTCAAGAATAACACAGCACACTGACCATTTCTGGAATCGGTCTCTGAGGATGAAGCC 981
Qy 1201 TACTACCAAGTGGTAGCAGAAACACAGCCGGGAACTGCTGTGTGCGCTGCGCCCTGCGG 1260
Db 982 ATTTATCAGTGTGTGGCCGAGAACAGTGGGGCTCATCACAGGCCAGTGGCCAGGCTGACC 1041
Qy 1261 GTAGTGGTGGCGGAGGGGCTGCCACAGCGCCCGGACTCGGGTCACAGCCACGCCCGCTGAGC 1320
Db 1042 GTACTGTGGCTGAGGGGCTCCCGGGCTCCCGGCAATGTGCGGGCAGTCTCTGTGTCT 1101
Qy 1321 AGCTCCTCTGTGCTGGTGGCTTGGGAGCGGCTGAGTTGCACAGCGAGCAAAATCAATTGGC 1380
Db 1102 TCCACTGAGGTGCGTGTGTCTGTGAGTGAGCGGCTGGCCAAACACCAAGGAGTCAATCGGC 1161
Qy 1381 TTCTCTCTTCACTACCAAGGCAAGGGGAGTGGACAAATGTGGAGTACCAAGTTTGCAGTA 1440
Db 1162 TACGTCCTGCACATCAGGAAGGCTGCTGACCCACCGGAGCTGGAGTATCAGGAGGCAGTC 1221
Qy 1441 AACATGACACACACAGACTGAGGTTGCGGACCTGGAAACCCAAACACGGATTATGAGTTC 1500
Db 1222 AGCAAGAGCACCTTTTCAGCACCTGGTCAAGCGACCTGGAGCCCTCCACAGCCTACAGTTTC 1281
Qy 1501 TACGTGGTGGCTACTCCAGCTGGGGCCAGCCGAACTCCAGCCAGCCCTGCTGTCAT 1560
Db 1282 TACATCAAGGCTTACACACCAAGGGGGCCAGCTCAGCCTCTGTGCCCCACCCCTAGTAGC 1341
Qy 1561 ACCTGGACGATGTCCCGAGCGCAGCACCCCGAG 1593
Db 1342 ACCCTGGGTGAAGCCCTGCCCCACCCCACTG 1374
```

Search completed: October 14, 2003, 10:17:18
Job time : 1181.58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 17:10:36 ; Search time 1668.34 Seconds
(without alignments)
9605.664 Million cell updates/sec

Title: US-09-754-997A-1
Perfect score: 6176
Sequence: 1 atggcgggggcgacacggg.....taataaactatttccagac 6176

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1731049 seqs, 1297405648 residues

Total number of hits satisfying chosen parameters: 3462098

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:**
1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgm2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgm2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*
4: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgm2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgm2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgm2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgm2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgm2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgm2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgm2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgm2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgm2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgm2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgm2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgm2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	6176	100.0	6176	10	US-09-754-997A-1
2	3756	60.8	3756	10	US-09-754-997A-45
3	2804.2	45.4	3753	12	US-10-239-663-29
4	2796	45.3	2796	10	US-09-754-997A-3
5	2681.8	43.4	3741	10	US-09-908-193-1
6	1522.4	24.7	2223	12	US-10-239-663-28
7	825	13.4	825	10	US-09-754-997A-5
8	316.6	5.1	2976	14	US-10-231-353-11
9	316.6	5.1	3210	14	US-10-231-353-1
10	316.6	5.1	3219	14	US-10-231-353-17
11	316.6	5.1	3874	14	US-10-231-353-31
12	315.4	5.1	3453	14	US-10-231-353-7
13	288	4.7	288	10	US-09-754-997A-19
14	252	4.1	252	10	US-09-754-997A-23
15	249	4.0	249	10	US-09-754-997A-15
16	249	4.0	249	10	US-09-754-997A-17

17	246	4.0	246	10	US-09-754-997A-21	Sequence 21, Appl
18	243	3.9	243	10	US-09-754-997A-7	Sequence 7, Appl
19	217	3.5	2481	14	US-10-231-353-15	Sequence 15, Appl
20	217	3.5	2715	14	US-10-231-353-5	Sequence 5, Appl
21	217	3.5	2724	14	US-10-231-353-19	Sequence 19, Appl
22	217	3.5	2958	14	US-10-231-353-9	Sequence 9, Appl
23	213.8	3.5	2139	14	US-10-231-353-21	Sequence 21, Appl
24	213.8	3.5	2382	14	US-10-231-353-27	Sequence 27, Appl
25	195	3.2	195	10	US-09-754-997A-13	Sequence 13, Appl
26	192	3.1	192	10	US-09-754-997A-9	Sequence 9, Appl
27	189	3.1	189	10	US-09-754-997A-11	Sequence 11, Appl
28	151.2	2.4	1875	14	US-10-231-353-23	Sequence 23, Appl
29	122.2	2.0	249	10	US-09-933-797-25	Sequence 25, Appl
30	119	1.9	909	14	US-10-231-353-13	Sequence 13, Appl
31	119	1.9	1143	14	US-10-231-353-3	Sequence 3, Appl
32	117	1.9	1644	14	US-10-231-353-25	Sequence 25, Appl
33	117	1.9	1887	14	US-10-231-353-29	Sequence 29, Appl
34	85.6	1.4	439	10	US-09-728-445-856	Sequence 856, App
35	71	1.1	4277	11	US-09-946-374-57	Sequence 57, Appl
36	71	1.1	4277	12	US-10-015-387A-57	Sequence 57, Appl
37	71	1.1	4277	12	US-10-006-130A-57	Sequence 57, Appl
38	71	1.1	4277	12	US-10-199-672-439	Sequence 439, App
39	71	1.1	4277	12	US-10-006-172A-57	Sequence 57, Appl
40	71	1.1	4277	12	US-10-187-749-439	Sequence 439, App
41	71	1.1	4277	12	US-10-194-457-439	Sequence 439, App
42	71	1.1	4277	12	US-10-184-642-439	Sequence 439, App
43	71	1.1	4277	12	US-10-196-747-439	Sequence 439, App
44	71	1.1	4277	12	US-10-015-392A-57	Sequence 57, Appl
45	71	1.1	4277	12	US-10-017-253A-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1

US-09-754-997A-1
; Sequence 1, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; FILE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6176
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3756)
US-09-754-997A-1

Query Match	100.0%	Score 6176;	DB 10;	Length 6176;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 6176;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	ATGGCGGGGGGACACACGGGGCGGGCTCTCTGCTGACCTTCTGCTGTCGCGG	60	
Db	1	ATGGCGGGGGGACACACGGGGCGGGCTCTCTGCTGTCGACCTTCTGCTGTCGCGG	60	
Oy	61	CGCGGGGAGCTGCATTGCCCGGAGACAACTGTCAAGCTGAGCTGTGATGAGGAGCC	120	
Db	61	CGCGGGGAGCTGCATTGCCCGGAGACAACTGTCAAGCTGAGCTGTGATGAGGAGCC	120	
Oy	121	CTGCAAGTGATCTCTGGGGCCCTGAGCAGGCTGTGGTGTGGACTGCACTTTGGGGCTACA	180	

Db	121	CTGCAAGTGATCTCGGGCTGAGCAGGCTGTGTGCTGCACTGCTTTGGGGCTACA	180	1201	TACTACCAAGTCGTAGCAGAAAAACAGCGCGGGAACTGCCTGTGTGCGCTGCGCCCCCTGGCG	1260
Qy	181	GCTGCTGGGCTCCGACCGAGGTGACATGGAGCAAGATGGAGACTGTACTAGAGCAT	240	1261	GTAGTGTGTGCGAGGGCTGCCAGCGCCCCGACTCGGGTCAAGACCAACGCGCGCTGAGC	1320
Db	181	GCTGCTGGGCTCCGACCGAGGTGACATGGAGCAAGATGGAGACTGTACTAGAGCAT	240	1261	GTAGTGTGTGCGAGGGCTGCCAGCGCCCCGACTCGGGTCAAGACCAACGCGCGCTGAGC	1320
Qy	241	GAGAACTTGACCTGCTACCAATGCTCCCTGTGTGCTGTCTCACCCTTAGAGCAAGAA	300	1321	AGCTCCTCTGTGTGTGCTGGAGCGGCTGTGATGTGCAACAGAGCAAAATCAATTTGGC	1380
Db	241	GAGAACTTGACCTGCTACCAATGCTCCCTGTGTGCTGTCTCACCCTTAGAGCAAGAA	300	1321	AGCTCCTCTGTGTGTGCTGGAGCGGCTGTGATGTGCAACAGAGCAAAATCAATTTGGC	1380
Qy	301	GACAGCAATGATGAGGAGCTCTTAGGATCTGGAAGTCACTGAGGGCAGCTATTCTCTGT	360	1381	TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGCAATGTGGAGTACAGATTTTCAGTGA	1440
Db	301	GACAGCAATGATGAGGAGCTCTTAGGATCTGGAAGTCACTGAGGGCAGCTATTCTCTGT	360	1381	TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGCAATGTGGAGTACAGATTTTCAGTGA	1440
Qy	361	CTGGCCCAACGCCCTAGAGTGTGGCCAGCCAGGTTGCTGTGTGCTTGAAGTTCACCA	420	1441	AACAATGACACACAGAGCTGAGGTTCGGGACCTTGGAACCCCAACCGATTTATCAGTTTC	1500
Db	361	CTGGCCCAACGCCCTAGAGTGTGGCCAGCCAGGTTGCTGTGTGCTTGAAGTTCACCA	420	1441	AACAATGACACACAGAGCTGAGGTTCGGGACCTTGGAACCCCAACCGATTTATCAGTTTC	1500
Qy	421	CTCGAAGACTTCTCTCTGCAACCCGAGTCCAGATTTGTGAGGAGAAACGGGACAGCACG	480	1501	TACGTGTGTGGCTACTCCAGCTGTGGGGCCAGCGGAACTTCAGCCAGCCCTGTGTGCAT	1560
Db	421	CTCGAAGACTTCTCTCTGCAACCCGAGTCCAGATTTGTGAGGAGAAACGGGACAGCACG	480	1501	TACGTGTGTGGCTACTCCAGCTGTGGGGCCAGCGGAACTTCAGCCAGCCCTGTGTGCAT	1560
Qy	481	TTTGAATGCCACCAAGGGCTTCCAGGCCCTCATATTACTTTGGGAAAGGACCAAGGTG	540	1561	ACACTGTGAGCATGTCTCCAGCTGCGGACACCCAGCTTACCTTTGTCAGCCCAACCCCTCG	1620
Db	481	TTTGAATGCCACCAAGGGCTTCCAGGCCCTCATATTACTTTGGGAAAGGACCAAGGTG	540	1561	ACACTGTGAGCATGTCTCCAGCTGCGGACACCCAGCTTACCTTTGTCAGCCCAACCCCTCG	1620
Qy	541	ACCGTGTGAGAGCCCGGCTCATCTCTTCCAGTGGCTCTCCAGTCCCTAGATCTTAGAT	600	1621	GACATCAGGTGGCATGTGCTGCTCCAGCTGAGCACTTGTGTGAGCAAGCTGTGAAG	1680
Db	541	ACCGTGTGAGAGCCCGGCTCATCTCTTCCAGTGGCTCTCCAGTCCCTAGATCTTAGAT	600	1621	GACATCAGGTGGCATGTGCTGCTCCAGCTGAGCACTTGTGTGAGCAAGCTGTGAAG	1680
Qy	601	GTCAGGACATGATCAGGCTCTACCGCTGCTGCGCACCAATTCAGCCCGCCAAACGA	660	1681	TACAAAGATAGATGACGTTTGGGAAAGGAAGATCAGGTTTTCTCCACAGGTGCTGGA	1740
Db	601	GTCAGGACATGATCAGGCTCTACCGCTGCTGCGCACCAATTCAGCCCGCCAAACGA	660	1681	TACAAAGATAGATGACGTTTGGGAAAGGAAGATCAGGTTTTCTCCACAGGTGCTGGA	1740
Qy	661	TTCAGCCAGAGGCTCTGCTCACTGTGGCTCAGAGGCTCTTTGGAGGCTACAGGGGG	720	1741	AATGAGACAACTTTACGTTAAACTCACTTACGCCAAACAAAGTGTACCGAGTCCGGATT	1800
Db	661	TTCAGCCAGAGGCTCTGCTCACTGTGGCTCAGAGGCTCTTTGGAGGCTACAGGGGG	720	1741	AATGAGACAACTTTACGTTAAACTCACTTACGCCAAACAAAGTGTACCGAGTCCGGATT	1800
Qy	721	CAGGATGTGTGCTATTGTGGAGCCCGAGAGAACACACCGTAGTGTCTGACAGATGTA	780	1801	TCAGCTGGCACTGGCGCTGCTTATGAGTCTTCTCAGTGTGATGTCAGCAGGACACCT	1860
Db	721	CAGGATGTGTGCTATTGTGGAGCCCGAGAGAACACACCGTAGTGTCTGACAGATGTA	780	1801	TCAGCTGGCACTGGCGCTGCTTATGAGTCTTCTCAGTGTGATGTCAGCAGGACACCT	1860
Qy	781	GTGATGAGTGTGCTGCTGCTGACCCACCCCTTTGTGTCTTGGTCCGACAGGAT	840	1861	GGTGTGCAACACAGAGCCATGTTCCCTTTGGCCCTGCAAGAAATGAAGGTGAGGGCAAG	1920
Db	781	GTGATGAGTGTGCTGCTGCTGACCCACCCCTTTGTGTCTTGGTCCGACAGGAT	840	1861	GGTGTGCAACACAGAGCCATGTTCCCTTTGGCCCTGCAAGAAATGAAGGTGAGGGCAAG	1920
Qy	841	GGAAAGCTATCTCCAGGATGTATGTTCTGGGCGGACCAATCTACTCATCGCCAGC	900	1921	ATGAGAGTCCCTGTGTGTGTCAGCGCGCCCTCACCCACCCAGATCTCTGTGATAC	1980
Db	841	GGAAAGCTATCTCCAGGATGTATGTTCTGGGCGGACCAATCTACTCATCGCCAGC	900	1921	ATGAGAGTCCCTGTGTGTGTCAGCGCGCCCTCACCCACCCAGATCTCTGTGATAC	1980
Qy	901	GCGAGGCTCTGAGTCTATGTCTGCGAGCCAAAGCCCTCTACGCGTGAC	960	1981	AAACTCTACTGGGAGAGGTGGAAACAGAGGAGGAGGAGATGTTGTAACCGCCCAAGG	2040
Db	901	GCGAGGCTCTGAGTCTATGTCTGCGAGCCAAAGCCCTCTACGCGTGAC	960	1981	AAACTCTACTGGGAGAGGTGGAAACAGAGGAGGAGGAGATGTTGTAACCGCCCAAGG	2040
Qy	961	TTCCGCACTCGGGCTGAGCTCCGAGTCTTGTGTCGCGGGCTCCGAGGACCC	1020	2041	GGTGTGAGAGTCAAGCTTTGGGACGTGCGGGCCCTGCGGCTGAAGAGAAAGTGAAGCAG	2100
Db	961	TTCCGCACTCGGGCTGAGCTCCGAGTCTTGTGTCGCGGGCTCCGAGGACCC	1020	2041	GGTGTGAGAGTCAAGCTTTGGGACGTGCGGGCCCTGCGGCTGAAGAGAAAGTGAAGCAG	2100
Qy	1021	GAGCGCTCTCGGGAGCGCGGGCCAGCACCGCGCTTGTGTGCGGGGCTCCGGGAG	1080	2101	TATGAATGACCCAGTTAGTCCCTGGCAGCGCGTACGAGGTGAGCTCGTAGCTTTCAAC	2160
Db	1021	GAGCGCTCTCGGGAGCGCGGGCCAGCACCGCGCTTGTGTGCGGGGCTCCGGGAG	1080	2101	TATGAATGACCCAGTTAGTCCCTGGCAGCGCGTACGAGGTGAGCTCGTAGCTTTCAAC	2160
Qy	1081	CCAGCGCCCGGCTGCTGCTGAGCGAGGATCCGTTGGGACCCAAATGGGCGGCTC	1140	2161	AAACACGAGGAGCGGCTTACGCTGTGTGTGAAAGGGCAAGACGAGAGGCGCCACGCCA	2220
Db	1081	CCAGCGCCCGGCTGCTGCTGAGCGAGGATCCGTTGGGACCCAAATGGGCGGCTC	1140	2161	AAACACGAGGAGCGGCTTACGCTGTGTGTGAAAGGGCAAGACGAGAGGCGCCACGCCA	2220
Qy	1141	AAGGTGAGGCGGCTGAGCTGCTGATCATCTCAGATCGGCTGAGGACGCTGGC	1200	2221	GACCTGCTTATCCAGAGGGGGCCACCGCTGCTCTCTGCCCATGTCCACGACAGTCAAAC	2280
Db	1141	AAGGTGAGGCGGCTGAGCTGCTGATCATCTCAGATCGGCTGAGGACGCTGGC	1200	2221	GACCTGCTTATCCAGAGGGGGCCACCGCTGCTCTCTGCCCATGTCCACGACAGTCAAAC	2280
Qy	1201	TACTACAGTGTGAGCAAAACAGCGCGGAACTGCTGTGCTGCGCCCTGGCG	1260	2281	AGCTCCACTTCCATTTGGCTTCGGTGGAAAGGCCAGACTTTTACCCTGTCAAGATTGTC	2340
				2281	AGCTCCACTTCCATTTGGCTTCGGTGGAAAGGCCAGACTTTTACCCTGTCAAGATTGTC	2340


```
Db 4501 |||||CTTGTCTTATGTCGTGTTCTCGTCCCTGAGTTCAACCTATGCACTCTTCTTAACA 4560
Qy 4561 ACATGACTACCTCATGTCGTCTTACACATAGTGTGACCCCTGGGTCCCAAGCTCCC 4620
Db 4561 ACATGACTACCTCATGTCGTCTTACACATAGTGTGACCCCTGGGTCCCAAGCTCCC 4620
Qy 4621 CTGCCAACCGCTTCTCTGGGAGAGTCCCACTCCCAAGTAGATCTGGAAGAGACCCCTG 4680
Db 4621 CTGCCAACCGCTTCTCTGGGAGAGTCCCACTCCCAAGTAGATCTGGAAGAGACCCCTG 4680
Qy 4681 TGGCTTGTCTGGCTGCCCTCCCTCTGGTGTGAGATGAGAAGTCTTCTATGGAAGAGAT 4740
Db 4681 TGGCTTGTCTGGCTGCCCTCCCTCTGGTGTGAGATGAGAAGTCTTCTATGGAAGAGAT 4740
Qy 4741 GAGTCCAGCTGCACAGGGGAGCCCAAGAGGGGTAGGGAGTGAAGAACAGAGGCTGA 4800
Db 4741 GAGTCCAGCTGCACAGGGGAGCCCAAGAGGGGTAGGGAGTGAAGAACAGAGGCTGA 4800
Qy 4801 AAAAAAATGGCTGCCACCCATCTGCACAGAGAGATGGGTGTGCTTTTGAAGTGCAGTC 4860
Db 4801 AAAAAAATGGCTGCCACCCATCTGCACAGAGAGATGGGTGTGCTTTTGAAGTGCAGTC 4860
Qy 4861 CTGGCTGAATCTGAAGGGGTGAGGAGGGAGCTACTGGGGCTGCCATGGCTCACTCC 4920
Db 4861 CTGGCTGAATCTGAAGGGGTGAGGAGGGAGCTACTGGGGCTGCCATGGCTCACTCC 4920
Qy 4921 CTGACCTGGAGCCCTGACCTGGCTTACAGTAGAGATGGGTGTGCTTTTGAAGTGCAGTC 4980
Db 4921 CTGACCTGGAGCCCTGACCTGGCTTACAGTAGAGATGGGTGTGCTTTTGAAGTGCAGTC 4980
Qy 4981 AAGGGAAGTCTTTGATAGAGAAAGGGGCTGGCTCATTTTATTTTATCTTTTATCA 5040
Db 4981 AAGGGAAGTCTTTGATAGAGAAAGGGGCTGGCTCATTTTATTTTATCTTTTATCA 5040
Qy 5041 CTGAATCCCAAAATCATCTTTACACAAAGGGGCAAGCTGAGTGTGCTTCTGAGTCAC 5100
Db 5041 CTGAATCCCAAAATCATCTTTACACAAAGGGGCAAGCTGAGTGTGCTTCTGAGTCAC 5100
Qy 5101 AAGAGCCATGCCATCTCTCTGGTTCTCAGCTCAGTCATGTCAGAGATGTCAGATCCA 5160
Db 5101 AAGAGCCATGCCATCTCTCTGGTTCTCAGCTCAGTCATGTCAGAGATGTCAGATCCA 5160
Qy 5161 GTGGCATCTGCTCTCTGTCGACATCTTTCTATTTTCAACTGGCTGCGACATCAAGTGT 5220
Db 5161 GTGGCATCTGCTCTCTGTCGACATCTTTCTATTTTCAACTGGCTGCGACATCAAGTGT 5220
Qy 5221 AACTCTGGCTTCTGGGCAAGTTAGAAATACCAAGTCTATTTTCCCTTTATTTTATTTA 5280
Db 5221 AACTCTGGCTTCTGGGCAAGTTAGAAATACCAAGTCTATTTTCCCTTTATTTTATTTA 5280
Qy 5281 TTTTATTTTATTTTATGTCCTTTCAGTGGAGTGTGAGCTTCTGAAGGCTGTGTTTAT 5340
Db 5281 TTTTATTTTATTTTATGTCCTTTCAGTGGAGTGTGAGCTTCTGAAGGCTGTGTTTAT 5340
Qy 5341 AGCTCTGTGTGCTCACTCATGTTTGAACCCCAAGCTTCTCTCTCTCTCTCTCTCAGC 5400
Db 5341 AGCTCTGTGTGCTCACTCATGTTTGAACCCCAAGCTTCTCTCTCTCTCTCTCTCAGC 5400
Qy 5401 CAGCTATGATAACACATTAAGATTTAATGCTGGCTTCTGATCTCAATTAAGACAGGAT 5460
Db 5401 CAGCTATGATAACACATTAAGATTTAATGCTGGCTTCTGATCTCAATTAAGACAGGAT 5460
Qy 5461 TGTCACTTGAACCTTCTATAGCAATTCAGAGTGGCCACCGGCAACACCGATGTTT 5520
Db 5461 TGTCACTTGAACCTTCTATAGCAATTCAGAGTGGCCACCGGCAACACCGATGTTT 5520
Qy 5521 CTTTCATGCTCTGAAGTCTCAAGGCTCATTTTGTGTTTCTGTTTCTGATTTCTTCTCC 5580
Db 5521 CTTTCATGCTCTGAAGTCTCAAGGCTCATTTTGTGTTTCTGTTTCTGATTTCTTCTCC 5580
Qy 5581 TTGCTTTGAATGAATTAACCGTTTAAACAGTAGGCTCTTAGCATCACACACATAGTCAT 5640
Db 5581 TTGCTTTGAATGAATTAACCGTTTAAACAGTAGGCTCTTAGCATCACACACATAGTCAT 5640
```

```
Db 5581 TTGCTTTGAATGAATTAACCGTTTAAACAGTAGGCTCTTAGCATCACACACATAGTCAT 5640
Qy 5641 TCCTCATGTCCTTCTGTTTAAACAGCACTTGAAGTCTCTGGTCTTAAATTAATAGCTGCAAA 5700
Db 5641 TCCTCATGTCCTTCTGTTTAAACAGCACTTGAAGTCTCTGGTCTTAAATTAATAGCTGCAAA 5700
Qy 5701 TGAGCAATTTATATACCCATTAGGCTGGGTGGAATTTGTTCTCAAAAGCAAAATAGTAA 5760
Db 5701 TGAGCAATTTATATACCCATTAGGCTGGGTGGAATTTGTTCTCAAAAGCAAAATAGTAA 5760
Qy 5761 TAAATCTGGTATCTGCCCTATTAACCTACAGTTGATTAAGAAAGTAGCCAGACTCCTAGCA 5820
Db 5761 TAAATCTGGTATCTGCCCTATTAACCTACAGTTGATTAAGAAAGTAGCCAGACTCCTAGCA 5820
Qy 5821 TTATATATGATTTGGGGTCTGAGTAACTGGGAGAGTGTAGCTTTTGTGACTTTGTAGCACC 5880
Db 5821 TTATATATGATTTGGGGTCTGAGTAACTGGGAGAGTGTAGCTTTTGTGACTTTGTAGCACC 5880
Qy 5881 AGGTCCTTATAGGAAGTCTGTGTGGCTTTTACAGGGCAATAGTCCCTTGTGCTTTGCC 5940
Db 5881 AGGTCCTTATAGGAAGTCTGTGTGGCTTTTACAGGGCAATAGTCCCTTGTGCTTTGCC 5940
Qy 5941 ATGGATGCTTAAAGTCTTTTGGAGTCTCATTTAAGAAATCTCTTCTCGAAGCATGACAA 6000
Db 5941 ATGGATGCTTAAAGTCTTTTGGAGTCTCATTTAAGAAATCTCTTCTCGAAGCATGACAA 6000
Qy 6001 GTGATTCGCAATCTTACATGCTCACTGTTTACCTGGCTAGTATTGTGCTGGGTATTT 6060
Db 6001 GTGATTCGCAATCTTACATGCTCACTGTTTACCTGGCTAGTATTGTGCTGGGTATTT 6060
Qy 6061 AATTGCACCTTCCAGCATCATGCTTCTCTTACAAATATGATATCTTTATTTGTATCAC 6120
Db 6061 AATTGCACCTTCCAGCATCATGCTTCTCTTACAAATATGATATCTTTATTTGTATCAC 6120
Qy 6121 TAAAGTGTGATCATGATCTCTCTCTGTAAGAAATTAATAACTATTTTTCAGAC 6176
Db 6121 TAAAGTGTGATCATGATCTCTCTCTGTAAGAAATTAATAACTATTTTTCAGAC 6176
```

RESULT 2

```
US-09-754-997A-45
; Sequence 45, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 3756
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-754-997A-45
```

```
Query Match 60.8%; Score 3756; DB 10; Length 3756;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ATGGCGCGGCGGCGACACGCGGCGGCTCTCTGGTGTCTGACCTTCTGCTGCTGCTCGCG 60
Db 1 ATGGCGCGGCGGCGACACGCGGCGGCTCTCTGGTGTCTGACCTTCTGCTGCTGCTCGCG 60
Qy 61 CGCGGGGAGCTGCTCATTTGCCCCCAGGAGACACCTGTCAAGCTGAGCTGTGATGAGGACCC 120
Db 61 CGCGGGGAGCTGCTCATTTGCCCCCAGGAGACACCTGTCAAGCTGAGCTGTGATGAGGACCC 120
```


QY 183 TGCTGGGCTCGACACAGGTGACATGAGCAAGATGAGAGACATGTACTAGAGCATGA 242
Db 189 CGCTGGACCCCCACACAGGTGACCTGGAGCAAGATGGGAGACACCTGCTGGAGCAAG 248
QY 243 GAACTGCACTGCTTACCAATGGCTCCCTGCTGGCTGTCTCACCCTAGAGCAAGA 302
Db 249 CCATTACCTGCTGCTCCCAATGGTTCCTGTGGCTGTCCAGCCACTAGACCCCAATGG 308
QY 303 CAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGTCACTAGAGGACGCTATTTCGTGCT 362
Db 309 CAG---TGACGAGTCAGTCCCTGAGGCTGTGGGGTCAATTGAAGCACTATTTCGTGCT 365
QY 363 GGCACACAGCCGCTAGAGATGGTGGCCAGCCAGAGTGTGCTGTGCTGAAGCTTGCACACT 422
Db 366 AGCCACAGCCGCTCGAGTGTGCGCAGCCAGACTGTCTGTCTGAAGCTTGCACACT 425
QY 423 CGAGACTTCTCTGCAACCCGAGTCCAGATTGTGGAGAGAACGGGACAGCAGCTT 482
Db 426 CGCAGACTTCTCTGCAACCCGAGTCTCAGACGCTGGAGAGAACGGGACAGCTCGCTT 485
QY 483 TGAATGCCACCAAGGCTTCCAGCCCTCATCTTACTTTGGGAAAGGACCAAGGTGAC 542
Db 486 TGAGTGCCACATTGAAGGCTGCGAGCTCCCATCTTACTTTGGGAAAGGACCAAGGTGAC 545
QY 543 CGTGCCTGAGAGCCCCGCTCATCACTTCCCAAGTGGCTCTCCAGATCTTAGATGT 602
Db 546 ATTGCTGAGGAGCTCGCTCATGTGCTTCCCAAGCGGCTCTTTCAGATCTCTGATGT 605
QY 603 CCAGACATGATGACAGGCTCTACCGCTGGTGGCCACCAATTCAGCCCGCCCAAGATT 662
Db 606 TCAGAGAGTGAATGAGGCGCTTACCGCTGGTGGCCACCAATTCAGCTCGCCAGCACTT 665
QY 663 CAGCCAGAGGCTCGCTCACTGTGGCTCAGAGGCTCTTTGAGGCTTACAGGGGCA 722
Db 666 CAGCCAGAGGCTCTCACTGTGGCTCAGAGGCTCTTTGAGGCTTACAGGGGCA 725
QY 723 GGATGTGTCTATGTGGCAGCCCAAGAGAACCAACGATGTGTCTGGAAGATGTAGT 782
Db 726 GGAAGTGTCTATGTGGCAGCCCAAGAGAACCAACGATGTGTCTGGAAGATGTAGT 785
QY 783 GATGGAGTGTGGCTCTGTGACCCCAAGCTTGTGTCTGGGTCCAGAGATGG 842
Db 786 GATGGAATGTGGCTCTGTGACCCCAAGCTTGTGTCTGGGTCCAGAGATGG 845
QY 843 AAAGCTATCTCCAGATGTCTGTTCTGGGCGGACCAATCTTACTCATCGCCAGGCT 902
Db 846 GAAAGCTATCTCCAGATGTCTGTTCTGGGCGGACCAATCTTACTCATCGCCAGGCT 905
QY 903 CGAGCTCGGCACTGTGAGTCTATGTCTGCGAGCCCAAGCCCTTCACTGCTGACTT 962
Db 906 CGAGCTCGGCACTGTGAGTCTATGTCTGCGGCGCCCAAGCCCTTCACTGCTGACTT 965
QY 963 CGCCACTGGGCTGTGAGTCTCGAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
Db 966 CGCCACTGAGCGCTGAGTCTCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1025
QY 1023 GGCCTCTCGGAGCGGGGCGAGCAGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
Db 1026 GGCCTCTCGGAGCGGGGCGAGCAGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1085
QY 1083 ACGGCCGCTGCACTGGCTGCAAGCGGATCCGTTGCGAGCCCAATGCGGCGCTGCTGCT 1142
Db 1086 CGGGCCAGCGCTGCTGCTGCAAGCGGATCCGTTGCGAGCCCAATGCGGCGCTGCTGCT 1145
QY 1143 GGTGAGGGCGGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
Db 1146 GGTGAGGGCGGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205
QY 1203 CTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
Db 1206 CTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265

QY 1263 AGTGTGGCGAGGGGCTGCCAGCGGCCCGACTGCGGTTCAGAGCCAGCGCTGAGCAG 1322
Db 1266 GGTGTGGCGAGGGGCTGCCAGCGGCCCGACTGCGGTTCAGAGCCAGCGCTGAGCAG 1325
QY 1323 CTCCTCTGTGTGTGTGGAGCGGCTGTGAGTTCAGAGCGAGCAAAATCATTTGGCTT 1382
Db 1326 CTCCTCTGTGTGTGTGGAGCGGCTGTGAGTTCAGAGCGAGCAAAATCATTTGGCTT 1385
QY 1383 CTCCTCTCACTTACCAAAAGGCAAGGGAGTGGACAATGTGTGAGTACCAAGTTTGCAGTAAA 1442
Db 1386 CTCCTCTCACTTACCAAAAGGCAAGGGAGTGGACAATGTGTGAGTACCAAGTTTGCAGTAAA 1445
QY 1443 CAATGACACACACAGAGCTGCGAGTTCGGGACCTGGGAAACCAACAGGATTTAGATTCTA 1502
Db 1446 CAACGACACACACAGAGCTTACAGGTTTCGGGACCTGGGAAACCAACAGGATTTAGATTCTA 1505
QY 1503 CGTGTGGCTTACTTCCAGCTTGGGGCCAGCGAACTTCCAGCCAGCCCTGCTGTCATAC 1562
Db 1506 CGTGTGGCTTACTTCCAGCTTGGGGCCAGCGAACTTCCAGCCAGCCCTGCTGTCATAC 1565
QY 1563 ACTGAGCATGTCTCCAGCGCAGCACTTACCTTGTCTCAGCCCTCAACCTCTCGGA 1622
Db 1566 ACTGAGCATGTCTCCAGCGCAGCACTTACCTTGTCTCAGCCCTCAACCTCTCGGA 1625
QY 1623 CATCAGGGTGGCATGGCTGCTGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTT 1682
Db 1626 CATCAGGGTGGCATGGCTGCTGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTT 1685
QY 1683 CAGATAGATGTAGCTTGGGAGGAGATCAGGTTTCTTCCAGCGAGGCTCTGAGAA 1742
Db 1686 CAGATAGATGTAGCTTGGGAGGAGATCAGGTTTCTTCTACTGAGGTGCGAGGAA 1745
QY 1743 TGAGACAACTTACGTTTAAACTCACTTTCAGCCAAACAAAGTGTACCGAGTCCGAGTTTC 1802
Db 1746 TGAGACAACTTACGTTTAAACTCACTTTCAGCCAAACAAAGTGTATCGAGTACCGATTC 1805
QY 1803 AGCTGGCATCTGGCTGCTGCTTATGAGTTCCTTTCAGTGTGATGAGCAAGCACTTGG 1862
Db 1806 GGTGTGTACAGCAGCGGCTTCCGGGCTTCCAGTGTGATGATCAGAGCACTTCCAG 1865
QY 1863 TGTGCAACACAGAGCATGTTCCCTTTCGCTGCGGCTGAGATTTGAGGTGAGGGCAAGAT 1922
Db 1866 TATGCAACACAGAGCATGTTCCCTTTCGCTGCGGCTGAGATTTGAGGTGAGGGCAAGAT 1925
QY 1923 GGAGTCTCTGTGTGTGATGAGCAGCGCTTCCAGCCAGATCTCTGATACAA 1982
Db 1926 GGAGTCTCTGTGTGTGATGAGCAGCGCTTCCAGCCAGATCTCTGATACAA 1985
QY 1983 ACTCTACTGGGAGAGTGGGAAACAGAGGAGGAGATGTTGATGATGATGATGATGATGAT 2042
Db 1986 ACTATAATGGCGGAGTGGGGCTGAGGAGGAGGCAATGGCGATGCTGCTGCGAGGG 2045
QY 2043 TCGTGGAGATCAAGCTTGGGACGTCGGGCTGCGGCTGAGAGGAGGAGGAGGAGGAGGAG 2102
Db 2046 CCGTGGAGACAGGCTTGGGATGTTGGGCTGCTGCGGCTCAAGAGGAGGAGGAGGAGGAG 2105
QY 2103 TGAATCAACAGTGTAGTCTGCTGAGCGCTGAGGTGAGGCTGCTGAGCTTTCACAA 2162
Db 2106 TGAGTCAACAGTGTAGTCTGCTGAGCGCTGAGGTGAGGCTGCTGAGCTTTCACAA 2165
QY 2163 ACAGAGAGCGCTACGCTGCTGTGTGGAAGGGCAGAGCGGAGGAGGCGCCACGCCAGA 2222
Db 2166 ACATAGGATGGCTATACGAGCAGTGTGGAAGGGCAGAGCGGAGGAGGCGCCACGCCAGA 2225
QY 2223 CCGTCTATCGAGGGGGGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2282
Db 2226 CATGCTATCGAGGGGGGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2285
QY 2283 CTCACATTCATTTGGCTTGGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2342
Db 2286 CTCACATTCATTTGGCTTGGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2345
QY 2343 CTACATGTGTGCTTCCGCGCTTCCGCGGCTCAGGAAATGCTTCTGCTGCTACCTACTATAC 2402

QY 364 GCCACAGCCCGTAGGAGTGTGGCCAGCCAGGTTGCTGTGCTCAAGCTTGCCACACTC 423
DB |||||
QY 301 GCCACAGCCCGTAGGAGTGTGGCCAGCCAGGTTGCTGTGCTCAAGCTTGCCACACTC 360
DB |||||
QY 424 GAAGACTTCTCTCTGACCCCGAGTCCACAGATTGTGGAGGAGAAAGGACACACGCTTT 483
DB |||||
QY 361 GAAGACTTCTCTCTGACCCCGAGTCCACAGATTGTGGAGGAGAAAGGACACACGCTTT 420
DB |||||
QY 484 GAATGCCACACCAAGGCGCTTCAGAGCCCGCATCATATTCTTGGGAAAGGAACAGGTGACC 543
DB |||||
QY 421 GAATGCCACACCAAGGCGCTTCAGAGCCCGCATCATATTCTTGGGAAAGGAACAGGTGACC 480
DB |||||
QY 544 GTGCCCTGAGGAGCCCGCGCTCATCTCTTCCCAAGTGGCTCCTCCAGATCTAGATGTC 603
DB |||||
QY 481 GTGCCCTGAGGAGCCCGCGCTCATCTCTTCCCAAGTGGCTCCTCCAGATCTAGATGTC 540
DB |||||
QY 604 CAGGACAGTGTAGTGCAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAACGATTC 663
DB |||||
QY 541 CAGGACAGTGTAGTGCAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAACGATTC 600
DB |||||
QY 664 AGCCAGAGGCGCTCGTCTCATCTGTGCGCTCAGAGGCTCTTGGAGGCTACCAAGGGGCGAG 723
DB |||||
QY 601 AGCCAGAGGCGCTCGTCTCATCTGTGCGCTCAGAGGCTCTTGGAGGCTACCAAGGGGCGAG 660
DB |||||
QY 724 GATGTGCTCATTTGTGGCAGCCCGCAGAGAACACACCGTAGTGTCTGGACAGATGGA 783
DB |||||
QY 661 GATGTGCTCATTTGTGGCAGCCCGCAGAGAACACACCGTAGTGTCTGGACAGATGGA 720
DB |||||
QY 784 ATGAGTGTGCTGGCTCTGTGACCCCGCTCTTGTGCTCTGCTGCTCGGAGGAGTGA 843
DB |||||
QY 721 ATGAGTGTGCTGGCTCTGTGACCCCGCTCTTGTGCTCTGCTGCTCGGAGGAGTGA 780
DB |||||
QY 844 AAGCTATCTCCACGAGTGTATGCTGTGCGCCGAGCAATCTACTCATGCCAGCGCG 903
DB |||||
QY 781 AAGCTATCTCCACGAGTGTATGCTGTGCGCCGAGCAATCTACTCATGCCAGCGCG 840
DB |||||
QY 904 CAGCTTCGCACTCTGAGTGTATGCTGTGCGGAGCAACAGCCCTCAAGGCTGACTTC 963
DB |||||
QY 841 CAGCTTCGCACTCTGAGTGTATGCTGTGCGGAGCAACAGCCCTCAAGGCTGACTTC 900
DB |||||
QY 964 GCACTGCGGCTGTGAGTCTCGAGTGTGCTGCGCCGAGCCATCTCGCAGGCAACCCGAG 1023
DB |||||
QY 901 GCACTGCGGCTGTGAGTCTCGAGTGTGCTGCGCCGAGCCATCTCGCAGGCAACCCGAG 960
DB |||||
QY 1024 GCGCTCTCGGAGCGCGGCGCAGCACCGCGGCTTGTGCTGCGGCGCTCGCGGAGGCA 1083
DB |||||
QY 961 GCGCTCTCGGAGCGCGGCGCAGCACCGCGGCTTGTGCTGCGGCGCTCGCGGAGGCA 1020
DB |||||
QY 1084 CGGCCCGCGCTGCACTGCGTGTGCAACCGGATCCCGTTGCGAGCCCAATGGGCGGCTCAAG 1143
DB |||||
QY 1021 CGGCCCGCGCTGCACTGCGTGTGCAACCGGATCCCGTTGCGAGCCCAATGGGCGGCTCAAG 1080
DB |||||
QY 1144 GTGCAGGCGGCTGGCGGAGCTTGTGCTCATCTCAGATCGGCTGCGAGGAGCTGGCTAC 1203
DB |||||
QY 1081 GTGCAGGCGGCTGGCGGAGCTTGTGCTCATCTCAGATCGGCTGCGAGGAGCTGGCTAC 1140
DB |||||
QY 1204 TACCAAGTGTAGCAGAGAAACAGCGCGGAGAACTGCTGTGCGCTGCGCCCTCGCGGCTA 1263
DB |||||
QY 1141 TACCAAGTGTAGCAGAGAAACAGCGCGGAGAACTGCTGTGCGCTGCGCCCTCGCGGCTA 1200
DB |||||
QY 1264 GTGFTGCGGAGGCGCTGCGGAGCCCGGCTCGGCTGCAAGCCACCGCCCTGAGCAGC 1323
DB |||||
QY 1201 GTGFTGCGGAGGCGCTGCGGAGCCCGGCTCGGCTGCAAGCCACCGCCCTGAGCAGC 1260
DB |||||
QY 1324 TCCTGTGCTGCTGGCTGCGGAGCGGCTGAGTTGCAAGGAGCAATCATTTGGCTTC 1383
DB |||||
QY 1261 TCCTGTGCTGCTGGCTGCGGAGCGGCTGAGTTGCAAGGAGCAATCATTTGGCTTC 1320
DB |||||
QY 1384 TCTCTTCACTACCAAGGAGGAGTGGAGCAATGTGGAGTACAGTTTGCAGTAAAC 1443
DB |||||
QY 1321 TCTCTTCACTACCAAGGAGGAGTGGAGCAATGTGGAGTACAGTTTGCAGTAAAC 1380
DB |||||
QY 1444 AATGACACACAGAGCTCGAGTTCGGGACCTGGGAACCCCAACAGGATTATGAGTTCTAC 1503
DB |||||

DB |||||
QY 1381 AATGACACACAGAGCTGAGGTTGCGGACCTGGGAACCCAAACGATTATGAGTTCTAC 1440
DB |||||
QY 1504 GTGFTGCGCTACTCTCCAGCTGGGGCCAGCCGCAACCTCCAGCCAGCCCTGTGTGATACA 1563
DB |||||
QY 1441 GTGFTGCGCTACTCTCCAGCTGGGGCCAGCCGCAACCTCCAGCCAGCCCTGTGTGATACA 1500
DB |||||
QY 1564 CTGAGCGATGTCTCCAGCGCAGCAGCCCGAGCTTACCTTGTTCAGCCCAACCCCTCGGAC 1623
DB |||||
QY 1501 CTGAGCGATGTCTCCAGCGCAGCAGCCCGAGCTTACCTTGTTCAGCCCAACCCCTCGGAC 1560
DB |||||
QY 1624 ATCAGGCTGGCATGCTGCTGCTCCCTCCAGCCCTGAGCAATGAGCAGGCTGAGGATAC 1683
DB |||||
QY 1561 ATCAGGCTGGCATGCTGCTGCTCCCTCCAGCCCTGAGCAATGAGCAGGCTGAGGATAC 1620
DB |||||
QY 1684 AAGATAGATGATCGGTTTGGGGAAGAGATCAGGTTTCTCCACCGAGGCTGCTGGAAT 1743
DB |||||
QY 1621 AAGATAGATGATCGGTTTGGGGAAGAGATCAGGTTTCTCCACCGAGGCTGCTGGAAT 1680
DB |||||
QY 1744 GAGACACAACTTACGTTAAACTCTCTCAGCCCAACAAAGTGTACCGAGTCCGATTTCA 1803
DB |||||
QY 1681 GAGACACAACTTACGTTAAACTCTCTCAGCCCAACAAAGTGTACCGAGTCCGATTTCA 1740
DB |||||
QY 1804 GCTGGCACTGGGCTGCTGCTATGAGTCCCTTCTCAGTGGATGCGAGCAGGACACTGGT 1863
DB |||||
QY 1741 GCTGGCACTGGGCTGCTGCTATGAGTCCCTTCTCAGTGGATGCGAGCAGGACACTGGT 1800
DB |||||
QY 1864 GTGCACAAACAGAGCAGTGTCTTCCCTTGGCCCTGAGAAATGGAAGTGGAGGCAAGATG 1923
DB |||||
QY 1801 GTGCACAAACAGAGCAGTGTCTTCCCTTGGCCCTGAGAAATGGAAGTGGAGGCAAGATG 1860
DB |||||
QY 1924 GAGTCCCTGCTGTGTCTATGAGCAGCCGCTCTCAGCCCAACCAAGTGTACCGAGTCCGATTTCA 1983
DB |||||
QY 1861 GAGTCCCTGCTGTGTCTATGAGCAGCCGCTCTCAGCCCAACCAAGTGTACCGATCTCTGGATACAA 1920
DB |||||
QY 1984 CTCTACTGGGAGAGTGGGAAACAGAGGAGGAGCAGATGCTGAGCCGCCCCCGAGGGGT 2043
DB |||||
QY 1921 CTCTACTGGGAGAGTGGGAAACAGAGGAGGAGCAGATGCTGAGCCGCCCCCGAGGGGT 1980
DB |||||
QY 2044 CGTGAGATCAAGCTTGGGACGCTCGGCGCTGCGGCTGGAAGAGAAAGTGAAGCAGTAT 2103
DB |||||
QY 1981 CGTGAGATCAAGCTTGGGACGCTCGGCGCTGCGGCTGGAAGAGAAAGTGAAGCAGTAT 2040
DB |||||
QY 2104 GAACTGACCCAGTGTAGTCCCTGGCAGCGCTGTAACGAGTGAAGCTGTAGCTTTCAACAA 2163
DB |||||
QY 2041 GAACTGACCCAGTGTAGTCCCTGGCAGCGCTGTAACGAGTGAAGCTGTAGCTTTCAACAA 2100
DB |||||
QY 2164 CACAGAGCGCTACGCTGCTGTGTGGAAGGCGCAGAGAGGCGCCCAACGCGAGC 2223
DB |||||
QY 2101 CACAGAGCGCTACGCTGCTGTGTGGAAGGCGCAGAGAGGCGCCCAACGCGAGC 2160
DB |||||
QY 2224 CTGCTATCCAGAGGGGCGCACCGCTGCTCTGCTCCCAATGTCAGCAGAGTCAACAGC 2283
DB |||||
QY 2161 CTGCTATCCAGAGGGGCGCACCGCTGCTCTGCTCCCAATGTCAGCAGAGTCAACAGC 2220
DB |||||
QY 2284 TCCACTTTCCATTTGGCTTGGTGAAGAGCAGACTTTTACCACTGTCAAGATTTGTCAAC 2343
DB |||||
QY 2221 TCCACTTTCCATTTGGCTTGGTGAAGAGCAGACTTTTACCACTGTCAAGATTTGTCAAC 2280
DB |||||
QY 2344 TACACTGTACGCTTGGGCGCTGAGGAGTGTCTTCCCTGGTCACTACTATACC 2403
DB |||||
QY 2281 TACACTGTACGCTTGGGCGCTGAGGAGTGTCTTCCCTGGTCACTACTATACC 2340
DB |||||
QY 2404 AGCTCTGGAGAGACATTTCTCATTTGGGCGCTGAAACCAATTTACCAAGTACGAGTTTGG 2463
DB |||||
QY 2341 AGCTCTGGAGAGACATTTCTCATTTGGGCGCTGAAACCAATTTACCAAGTACGAGTTTGG 2400
DB |||||
QY 2464 GTACAGTCCCAACGAGTGGATATGATGGGCGCTTTTGGCTCGCTGTAAGAGCTGCCACC 2523
DB |||||
QY 2401 GTACAGTCCCAACGAGTGGATATGATGGGCGCTTTTGGCTCGCTGTAAGAGCTGCCACC 2460
DB |||||
QY 2524 CTGCTGACCGGCTTCAACACCTCTCTGTGACCTGCGGCTGAGCCCTCGACACCAATCC 2583
DB |||||

Db 1143 CAAGTCCAGGCGCGGTGGCAGGCTGGTCAATCAACAGATGCGGCTGCAGGAGCGCGG 1202
Qy 1200 CTACTACAGTGGCTAGCAGAAACAGCGCGGGAATCTGCTGCGGCTGCGCCCTGCG 1259
Db 1203 CTACTACAGTGGCTAGCAGAAACAGCGCGGGAATGGGCTGCGCTGCGGCTGCGTGC 1262
Qy 1260 GGTAGTGGTGGCGGAGGGGCTGCCAGGCGCCCGGACTCGGGTCAACAGCAGCGCGCTGAG 1319
Db 1263 CGTGGTGGTGGCGGAGGGGCTGCCAGGCGCCCGGACTCGGGTCACTGCTAGCGCACTGAG 1322
Qy 1320 CAGCTCCTCTGCTGCTGGCTGGAGGGGCTGAGTTGCAACAGCGAGGCAATCATTTGG 1379
Db 1323 CAGCTCCGCTGCTGCTGGCTGGAGGGGCTGAGTTGCAACAGCGAGGCAATCATCGG 1382
Qy 1380 CTTCTCTCTTCACTACCAAGGCAAGGGAGTGCAATGTGGAGTACCAAGTTTGGAGT 1439
Db 1383 CTTCTCTCTTCACTACCAAGGCAAGGGAGTGCAATGTGGAGTACCAAGTTTGGAGT 1442
Qy 1440 AAACAATGACCAACAGAGCTGACAGTTTGGGACCTGGAACCCAAACAGGATTAAGTT 1499
Db 1443 GAAACAAGACCAACAGAACTACAGTTTGGGACCTGGAACCCAAACAGATTAAGTT 1502
Qy 1500 CTAGTGGTGGCTTACTCCAGCTGGGGCCAGCGCAACCTCCAGCGCCAGCGCTGGTGA 1559
Db 1503 CTAGTGGTGGCTTACTCCAGCTGGGAGCCAGCGCAACCTCCAGCGCCAGCGCTGGTGA 1562
Qy 1560 TACATGAGAGATGCTCCAGCGAGCACCCAGCTTACCTTGTCCAGCGCCCAACCCCTC 1619
Db 1563 CACATGAGATGATGCTCCAGGAGCACCCAGCTTCCCTGTCCAGCGCCCAACCCCTTC 1622
Qy 1620 GGACATCAGGTTGGCATGGCTGCCCTGCTCCAGCTGAGCAATGACAGTGGCTGAA 1679
Db 1623 GGACATCAGGTTGGCTGGCTGCCCTGCTCCAGCTGAGCAATGAGGAGGTTGGA 1682
Qy 1680 GTAAGAATAGAGTACGGTTTGGGAAAGAAAGATCAGGTTTCTCAACGAGGTCCTGG 1739
Db 1683 GTACAAAGTAGAATACGGTTTGGGAAAGAAAGATCAGATTTTCTCTAGAGTGGCAGG 1742
Qy 1740 AAATGAGACACAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1799
Db 1743 AAATGAGACACAGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1802
Qy 1800 TTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1859
Db 1803 TTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1862
Qy 1860 TGGTGTGCAACCAAGAGCCATGTTCCCTTTGCCCCCTGCAAGATTAAGAGTGAAGGCAAA 1919
Db 1863 CAGTATGCAACCAAGAGCCATGTTCCCTTTGCCCCCTGCAAGATTAAGAGTGAAGGCAAA 1922
Qy 1920 GATGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1979
Db 1923 GATGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1982
Qy 1980 CAAACTCTACTGGGAGAGTGGGAAACAGAGGAGGAGGAGTGGTGAACGCGCCCGCCAGG 2039
Db 1983 CAAACTATATTGGCGGAGTGGGAAACAGAGGAGGAGGAGTGGTGAACGCGCGCTGCGCAGG 2042
Qy 2040 GGGTGTGAGATCAAGCTTGGGAGCTGGGCGCGCTGGGCTGAGAGGAGGAGGAGGAGGAGG 2099
Db 2043 GGGCGTGGAGACCAAGCTTGGGATGTTGGGCGCTGCTGGCTCAAGAGAAAGTGAAGCA 2102
Qy 2100 GTATGAATGACCAAGTTAGTCCCTGGCAGGCGCTGACAGGTTGAAGCTCTGAGCTTTCAA 2159
Db 2103 GTATGAGTGAACCAAGTTAGTCCCTGGCAGGCGCTGACAGGTTGAAGCTCTGAGGCTTTCAA 2162
Qy 2160 CAAACAGAGGAGGCTAGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2219
Db 2163 CAAACATGAGGATGCTATGACAGAGTGTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2222
Qy 2220 AGACCTGCTATTCAGAGGGGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2279

Db 2223 AGACATGCTTATCCAGAGGGAGCCACCCCTGCTCCAGCCCACTGCTCATGCGGAATCAA 2282
Qy 2280 CAGCTCCACATTTCCATTTGGGCTTTCGGTGAAGAGCCAGAGCTTTTACCCTGCTCAAGATTGT 2339
Db 2283 CAGCTCCACATCCATCTGGCTTTCGGTGAAGAGCCAGATTTTCAACAGTCAAGATTGT 2342
Qy 2340 CAACTACATGTTAGCTTTCGGCCCTTCGGGGCTCAGGAATGCTTCCCTGCTGCTCACTA 2399
Db 2343 CAACTACATGTTAGCTTTCAGCCCTTCGGGGCTCAGGAATGCTTCCCTGCTGCTCACTA 2402
Qy 2400 TACAGCTCTGGGAGAGCAATTTCTCATTTGGGGCTGGAACCAATTTACCAAGTACGATT 2459
Db 2403 ---CAGTTTGGAGAGACATCTCATTTGGGGCTTGAAGCCATTTCAACCAATACGATT 2459
Qy 2460 TGGCGTACAGTCCCAACGAGTGGATATGATGGGCTTTCGGCTTTCGGTTCCTGCTAGAACGCTC 2519
Db 2460 TGGAGTGCAGTCTCAGCGCTGAGATGATGGGCTTTCGGCTTTCGGCTTTCGGTTCGGAGCGCTC 2519
Qy 2520 CACCTGCTGAGCGGCTTCAACACCTCTTTCGACTGCGCTGAGCGCCCTGAGCACCC 2579
Db 2520 CACCTGCTGAGCGGCTTCAACACCTCTTTCGACTGCGCTGAGCGCCCTGAGCACCC 2579
Qy 2580 ATCCACGTTGGTTTACATGCTGGTTCCTCCCAACGAGCGCCCAATGGTGAATTTGGAGTA 2639
Db 2580 GTCCACGTTGGGCTGCACTGGTGGCTTTCGGTGGCTTTCGGTGGGAGATTCGTGGAGTA 2639
Qy 2640 TCTAATTTCTTACAGCAACCAACACACAGCGCCGAAACACAGTGGGACCTGCTCACCAC 2699
Db 2640 TCTGATCTCTGTAAGAGAGCAACCAACAGCGCTTTCGGTGGGAGATTCGTGGAGTA 2699
Qy 2700 -----AGAGGAGAACTTTCAGTGCAGAGGCTTCATGGCTTACAGAGTGAACCTCGGTA 2753
Db 2700 GCAGGTTGAGGAGAACTTTCAGTGCAGAGGCTTCATGGCTTTCAGAGGAGGAGTTCGGTA 2759
Qy 2754 TTTCTTCAAGATGGGAGCGCCGCAACAGAGTGGGCTTTCGGGCTTTCGGCTTTCGGCTTTCGG 2813
Db 2760 CTTCTTCAAGATGGGAGCGCCGCAACAGAGTGGGAGCTTTCGGGCTTTCGGCTTTCGGCTTTCGG 2819
Qy 2814 TGTGATTTCTTCAAGAGCAATTTCTCAGTGCAGAGGCTTCATGGCTTACAGAGTGAACCTCGGTA 2873
Db 2820 TGTGATCAGCTTCCAGGAGAGCTGTCAGAGTTCGCTGCAATGCACTCAGTCAAGGAGTAT 2879
Qy 2874 CATGCTGGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 2933
Db 2880 CATGCTGGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 2939
Qy 2934 ACAAAGTCTCCACAGGAGCGCTTTCGGGATGCTTTCCTCAGAGGAGCCCGGAGAAACCC 2993
Db 2940 CCGCAGCGCCCAACAGGAGATCCCTTCCAGGCTTTCCTCCACCGCCACCGCCCGGAGATCC 2999
Qy 2994 AGCGCTTACACAGAGCTGGCTTGGGCTTCCAGTGTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 3053
Db 3000 CGCGCTGTATCCAGAGCTGGCTTGGCGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCT 3059
Qy 3054 GTCCCTGCTGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 3113
Db 3060 GTCCCTTGTGCAACCGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 3119
Qy 3114 GGCTGAAGTACACAGCTTATGAGTGGGAGTGTTCAGATTTTCAGATTTTCAGATTTTCAGATTTTCAG 3173
Db 3120 GGCTGAAGTACACAGCTTATGAGTGGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCT 3179
Qy 3174 AAAGATCTCTGGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3233
Db 3180 AAAGGT---AAGTGTCTTCAACCAAGCGGCTGAGCTGGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 3236
Qy 3234 GCTGCCCGCAGGAGTGTCTTCAAGCGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGG 3293
Db 3237 GCTGCCCGCAGG---CAGGCGCCCGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGG 3293
Qy 3294 AACCGGCGCAGACATCTGCTGCTGCAAGCGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGG 3353
Db 3294 AACCGGCGCAGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3347

QY 3354 AAAGAGCCGCTCCAGCCTGCGAGGAATCAGGTGGAAGCTGAGGTCAATGTCTCACTCCGA 3413
Db 3348 GAAGAGTCAACCCAGCCTGCGAGAACAGGTGGAAGCTGAAGTCAATGTCTCACTCTGA 3407
QY 3414 CTTGGGTGATCCAAAGATGTCTGACCTCCACCTCCAAAGACTTGAGAGCCAGAGAAC 3473
Db 3408 CTTTAGTGATCTAAACGGGAACCTGACCTCCATCTCCAAAGACTTGAGAGCCAGAGAAC 3467
QY 3474 ACTGACTGACAGAGCTCTGCTTCCAGCTCTGGAGCTGTGGATCTGTCTCAAGGACGA 3533
Db 3468 CTTGCTCCAGAGGCTCTGATCTCATCTCGGGTGTGGGATCTCAAGGAGGGGAGC 3527
QY 3534 CTGGCTGGCAGGAGCTGGAGGGTGCACCAACAAACAGTGGGCCAGAGAGGCTCAC 3593
Db 3528 CTGGCTGGCAGGAGCTGGAGGGTGTGAGCTGGCAGCCCCGGGCCAGAGACTTAC 3587
QY 3594 CTGCTTCCAGAGAGCCAGTGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 3653
Db 3588 CTGCTTCCAGAGAGCCAGTGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 3647
QY 3654 TATAGAGAGCCCTGGAAAGCTGCCAGCCCAAGCCCTGTGCTCTTAACAGTCTAG 3713
Db 3648 GCTAGAGAGAGCCCTGGAGATAGCTGCCAGCTCAAAATCCCTGCTCTGCTCTGCTCT 3707
QY 3714 CCCAAGCTTCCAGAGGCCCTGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 3747
Db 3708 CCCAAGCTTCCAGAGGCCCTGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTCT 3741

RESULT 6

US-10-239-663-28
; Sequence 28, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, P.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-239-663-28

Query Match 24.7%; Score 1522.4; DB 12; Length 2223;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 1786; Conservative 0; Mismatches 321; Indels 51; Gaps 2;

QY 64 GGGGAGCTGCTTGGCCCGCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGGACCCCTG 123
Db 76 GGGGAGCTGCTTGGCCCGCCAGGAGACAACTGTGAGCTGAGCTGTGAGTGGGGCCACTG 135
QY 124 CAAGTATCTTGGGCTTGGAGCAGGCTGTGTGCTGAGCTGCACTTTGGGGGCTACAGCT 183
Db 136 CAAGTATCTTGGGCTTGGAGCAGGCTGTGAGCTGCACTTTGGGGGCTGCTGCC 195

QY 184 GCTGGGCCCTCCGACCAAGGATGAGAGCACTGTGTAGAGCATGAG 243
Db 196 GCTGACACCCCAACAGGCTGACCTGAGCAAGGATGAGGACACCTCTGTGAGCAGCAG 255
QY 244 AACCTGACCTGCTACCCCAATGGCTCCCTGTGGCTGTCTCACCCCTTAGAGCAAGAGAC 303
Db 256 CACTTACACCTGCTGCCCAATGGTTCCTGTGGCTGTCTCCAGCCACTAGACACCCCAATGCG 315
QY 304 AGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTAGAGGAGCAGCTATTCTGTCTG 363
Db 316 AG---TGACGATCAGTCCCTGAGGCTGTGGGGTCAATTGAAGGCAACTATTCTGTGCTTA 372
QY 364 GCCCAGACCCCTTAGGAGTGTGTCAGCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 423
Db 373 GCCCAGACCCCTTAGGAGTGTGTCAGCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 432
QY 424 GAAGACTTCTCTCTGCAACCCAGGCTCCAGATTTGAGGAGGAGAACGGGACAGCAGCTTT 483
Db 433 GAGAGCTTCTCTCTGCAACCCAGGCTCCAGATTTGAGGAGGAGAACGGGACAGCAGCTTT 492
QY 484 GAATGCCACACCAAGGGCTCTCCAGCCCTCCATTTATCTTGGGAAAGAGCAGGCTGACC 543
Db 493 GAGTGCACATTTGAAGGGCTGCGAGCTCCCATCATTTACTTGGGAGAGGACAGGCTGACA 552
QY 544 GTGCTTGAAGAGCCCGGCTCATCATCTTTCCAAAGTGTCTCTCAGATCTCTAGATGTC 603
Db 553 TTGCTTGAAGAGCTTCCGCTCATCGTGTCTTCCAAACGGCTCTTTCAGATCTCTGATGTT 612
QY 604 CAGGACGATGATGAGGCTCTTACCGCTGGTGGCCACCAATTTCAAGCCCGCCAAACGATTC 663
Db 613 CAGGAGATGATGAGGCTCTTACCGCTGGTGGCCACCAATTTCAAGCTCGCCAGACATTC 672
QY 664 AGCCAGAGGCTCTGCTCATCTGTGGCCCTCAGAGGCTCTTTGAGAGCTTACCAAGGGGAG 723
Db 673 AGCCAGAGGCTCTTACTAGTGTGGCCCAAGAGGCTCTCTGTGGCTTCCAGAGGGGAG 732
QY 724 GATGTGTCAATTTGTGAGCCCGCAGAGAACACACAGTGTGTGTGTGTGTGTGTGTGTGT 783
Db 733 GAGCTGTCAATTTGTGAGCCCGCAGAGAACACACAGTGTGTGTGTGTGTGTGTGTGTGT 792
QY 784 ATGAGTGTGTGGCTCTGCTGACCCCACTTTTGTGTCTTGGTCTCGACAGATGGA 843
Db 793 ATGGAATGTGTGGCTCTGCTGACCCCACTTTTGTGTCTTGGTCTCGACAGAGCGG 852
QY 844 AAGCCTATCTCCAGGATGTCTATGCTGTGGGCGGACCAATCTACTCATCGCAGCGCG 903
Db 853 AAGCCCATCTCCAGATGTCTATGCTGTGGGCGGACCACTCTATTAATTTGCAAGCGG 912
QY 904 CAGCTTGGCACTCTGAGTCTATGCTGTGGGCGGACCAAGCCCCCTCAAGCTGTGCTTC 963
Db 913 CAGCTTGGCACTCTGAGTCTATGCTGTGGGCGGACCAAGCCCCCTCAAGCTGTGCTTC 972
QY 964 GCACTTGGGCTGTGAGCTTCCAGTGTGTGTGGCCCGCAGCCATCTCGGAGGACCCGAG 1023
Db 973 GCACTTGGAGCGCTGAGCTTCCGCTGTGTGTGGGCGCTCCCGCCATCTCAAGGGCGCCGAG 1032
QY 1024 GCGCTTCCGAGCGGGCCAGACCGCGCTGTGTGTGGCGGCTCTCGGGGAGCCCA 1083
Db 1033 GCGCTTCCGAGCGGGCCAGACCGCGCTGTGTGTGGCGGCTCTCGGGGAGCCG 1092
QY 1084 CCGCCCGCTGCACTGGCTGCAAGCGGGATCCCGTTGCGACCCCAATGGGCGCTCAAG 1143
Db 1093 CCGCCAGCGCTGCGCTGCTGCAACAGCGGGCGCGCTGTGGGCCCAAGCGGCGCTCAAG 1152
QY 1144 GTGAGGGCGGTGGCGGAGCTTGTGATCATCTCAGATGCGCTGTGAGGAGCGTGTGCTAC 1203
Db 1153 GTGAGGGCGGTGGCGGAGCTTGTGATCATCTCAGATGCGCTGTGAGGAGCGCTGTGCTAC 1212
QY 1204 TACCACTGTGAGCAGAAACAGCGCGGAACTGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 1263
Db 1213 TACCACTGTGAGCAGCGCGGAACTGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 1272
QY 1264 GTGCTGCGGAGGGGCTGCTCCAGCGCCCGGCTCGGCTCAGAGCCACGCGCTGAGCAGC 1323

```
Db 1273 GTGGTCCGAGGGGTGCTCCAGCGCCCGCCAGCGGGTCACTGCTACGCCACTGAGCAGC 1332
Qy 1324 TCCTCTGTGCTGTGCTGCTGGAGCGGCTGAGTTGCACAGGAGCAAAATCAATGGCTTC 1383
Db 1333 TCGCTGTGTGTGCTGCTGGAGCGGCGCGAGATGACAGCGAGCAGATCATCGGCTTC 1392
Qy 1384 TCTCTTCACTACCAAAAGGGAAGGGAGGTGGAATGTGGAGTACAGATTGTGAGTAAC 1443
Db 1393 TCTCTCACTACCAAGAGCAGCGGGCATGGCAATGTGGAATACCAAGTTTGCAGTGAAC 1452
Qy 1444 AATGACACACAGAGCTGCGAGTTTGGGACCTGGGAACCAACACGAGTATCAGTTCTAC 1503
Db 1453 AAGCACACACAGAACTACAGTTTGGGACCTGGGAACCAACACAGTATCAGTTCTAC 1512
Qy 1504 GTGGTGGCTTACTCCAGCTGGGGGCCAGCGCAACCTCCAGCCAGCCCTGGTGCATACA 1563
Db 1513 GTGGTGGCTTACTCCAGCTGGGAGCCAGCGCACCTCCACCCAGCAGTGGTGACACA 1572
Qy 1564 CTGGAGCATGTCCCGAGCGAGCACCCAGCTTACCTTGTCCAGCCCCAACCCCTCGGAC 1623
Db 1573 CTGGATGATGTAGGGCCCTTGAATCTCGAGTGGG----- 1608
Qy 1624 ATCAGGGTGCATGGCTGCCCTCCAGCTGAGCAATGGACAGGTGCTGAAGTAC 1683
Db 1609 -----AGCTTGGGCTTGGAGCAATGGGAGGTGCTGAAGTAC 1644
Qy 1684 AAGATAGAGTACGGTTTGGGGAAGAAAGATCAGGTTTCTCCACCGAGTGCCTGGAAT 1743
Db 1645 AAGATAGATACGGTTTGGGGAAGGAAGATCAGATTTCTTCTACTGAGTGGCGGGAAT 1704
Qy 1744 GAGACAAACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGGATTCA 1803
Db 1705 GAGACACAGCTTATGCTGAACCTCGCTTCAGCCAAACAAAGTGTATCGATGACGATTCG 1764
Qy 1804 GCTGGACATGGCGCTGGCTATGAGTTCCTTCTCAGTGCATGACGACAGGACACTGCT 1863
Db 1765 GCTGATACAGCAGCGGCTTCGGGGGCCCTCCAGTGGATGCATCAGAGGACGCCAGT 1824
Qy 1864 GTGCACAAACAGAGCCTATGCTCCCTTGGCCCTGAGAAATGGAAGTGAAGGCAAGATG 1923
Db 1825 ATGCACAAACAGAGCCTATGCTCCCTTGGCCCTGAGAGTTGAAGTGCAGGCAAGATG 1884
Qy 1924 GAGTCCCTGGTGTATGTCAGCGAGCGCCCTTCAACCCACCCAGATCTCTGGATACAA 1983
Db 1885 GAGTCCCTGGTGTATGTCAGCGAGCCACCCCTTCAACCCACCCAGATCTCTGGTACAA 1944
Qy 1984 CTCTACTGGGAGAGTGGGAACAGAGGAGGAGGAGATGTGACCGCCCCCAGGGGT 2043
Db 1945 CTATATTGGCGGAGGTGGGGCTGAGGAGGAGGCAATGGCGATCGCTGCAGGGGC 2004
Qy 2044 CGTGAGATCAAGCTTGGGACGTTCGGGCCCTGCGGCTGAAGAAAGTGAAGCAGTAT 2103
Db 2005 CGTGAGACAGCTTGGATGTGGGGCTGTGGGGCTGTCCGGCTCAAGAAAGTGAAGCAGTAT 2064
Qy 2104 GAACTGACCCAGTATGCTCTGGCAGCGCGTGTACGAGTGAAGCTGTAGCTTTCAACAA 2163
Db 2065 GAGTGACCCAGCTAGTGCCTGGCGCGGTGTACGAGTGAAGCTGTGGCTTTCAACAA 2124
Qy 2164 CAGGAGGAGGCTACGCTGTGTGGAAAGGGGAGAGGAGGAGGAGGAGGAGGAGGAG 2221
Db 2125 CATGAGGATGGCTATGACAGAGTGTGGAAGGGGAGAGGAGGAGGAGGAGGAGGAGG 2182
```

RESULT 7

```
US-09-754-997A-5
; Sequence 5, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
```

```
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Mus musculus
; PEPTIDE:
; NAME/KEY: CDS
; LOCATION: (1)...(825)
; US-09-754-997A-5
```

```
Query Match
Best Local Similarity 13.4%; Score 825; DB 10; Length 825;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2932 CGACAAAGCTCCACAGGGAAGCCCTTCCCGGATTTGCTCTCAGGCAACCCACAGAAAC 2991
Db 1 CGACAAAGCTCCACAGGGAAGCCCTTCCCGGATTTGCTCTCAGGCAACCCACAGAAAC 60

Qy 2992 CCAGCGCTCTACACAAAGAGCTCGGCTTGGGCTTCCAGTGTCTCTGTGCCCCATGAGTTG 3051
Db 61 CCAGCGCTCTACACAAAGAGCTCGGCTTGGGCTTCCAGTGTCTCTGTGCCCCATGAGTTG 120

Qy 3052 GAGTCCCTCGTGCATCTCTGCTCCCGAGGATTTGGTTCCTCCACCACTCTCAGATGTGAAGC 3111
Db 121 GAGTCCCTCGTGCATCTCTGCTCCCGAGGATTTGGTTCCTCCACCACTCTCAGATGTGAAGC 180

Qy 3112 AAGGCTGAAGTACACAGGCTTATGGGTGGCAGTGTTCAGAGTATGCGGGCCACTCCCAAG 3171
Db 181 AAGGCTGAAGTACACAGGCTTATGGGTGGCAGTGTTCAGAGTATGCGGGCCACTCCCAAG 240

Qy 3172 AGAAAGATCTCTGGGCTCAGGCGGGGACAAACTGGGCGAGGCTCTCGGCGAGGCTGT 3231
Db 241 AGAAAGATCTCTGGGCTCAGGCGGGGACAAACTGGGCGAGGCTCTCGGCGAGGCTGT 300

Qy 3232 GAGCTGCCCCAGGCTAGTGGTCCAGGCGGGCTCTGACCCGCTCTCTGCTGCTTCAGG 3291
Db 301 GAGCTGCCCCAGGCTAGTGGTCCAGGCGGGCTCTGACCCGCTCTCTGCTGCTTCAGG 360

Qy 3292 GGAACCGGCGCAGACACTCTCTGCTGCAAGCCCTGGTGTATGACGATATAAGAGCAACGG 3351
Db 361 GGAACCGGCGCAGACACTCTCTGCTGCAAGCCCTGGTGTATGACGATATAAGAGCAACGG 420

Qy 3352 AGAAAGAGCCCTGCCAGCCTGCAAGGAATCAGGTGGAAGCTGAGGTCAITTTCCATCC 3411
Db 421 AGAAAGAGCCCTGCCAGCCTGCAAGGAATCAGGTGGAAGCTGAGGTCAITTTCCATCC 480

Qy 3412 GACTTCGGTGCATCAAAAGGATGCTGACCTCCACCTCCAAAGCTGAGGCGAGGAA 3471
Db 481 GACTTCGGTGCATCAAAAGGATGCTGACCTCCACCTCCAAAGCTGAGGCGAGGAA 540

Qy 3472 CCAGCTGACTGACAGACACTCTGCTTCCACCTCTGAGGCTGTGGATCTGTCTCAAGAGCA 3531
Db 541 CCAGCTGACTGACAGACACTCTGCTTCCACCTCTGAGGCTGTGGATCTGTCTCAAGAGCA 600

Qy 3532 GACTGGCTGGGCGAGGAGCTGGGAGGGTGCCAAACCAACAGTGGGGCGAGAGGCTC 3591
Db 601 GACTGGCTGGGCGAGGAGCTGGGAGGGTGCCAAACCAACAGTGGGGCGAGAGGCTC 660

Qy 3592 ACCTGCTTGCAGAAAGCAGCAGTGGCTTCTGCTCTCTGCTCTCAGACTCCAGGCCAGCT 3651
Db 661 ACCTGCTTGCAGAAAGCAGCAGTGGCTTCTGCTCTCTGCTCTCAGACTCCAGGCCAGCT 720

Qy 3652 GCTATAGAGGAGGCGCCCTGGGGAAGAGCTGCCAGGCCAAGCCCTGTGTCTTACAGTC 3711
Db 721 GCTATAGAGGAGGCGCCCTGGGGAAGAGCTGCCAGGCCAAGCCCTGTGTCTTACAGTC 780
```



```

QY 1974 TGGATACAACTCTACTGGGAGAGGTGGGAACAGAGGAGGAGGAGAGTGTGACCGCCC 2033
Db 1713 GGGCTACAAAGCTGTACTACAGGAAGAAGGGCAGCAGAGAA----- 1754
QY 2034 CCCAGGGGCTCGTGGAGATCAAGCTTGGACCTGGGACCTGGGCCCCGTGGGCTGAAGAGAAAGT 2093
Db 1755 -----TGGGCCCAATTTCTTGGATACCAAGGA 1781
QY 2094 GAAGCAGTATGAATGACCCAGTTAGTCCCTGGCAGGCCGTACGAGGTGAAGCTGTGAGC 2153
Db 1782 CCTACTCTATCTCTCAGTGGCTTAGACCCCAAGAAATAATCATGTGAGACTCTCTGC 1841
QY 2154 TTTCAACAAACACGAGGACGGCTAGCTGTGTGTGGAGGGCAAGACGGAGAGGGCCC 2213
Db 1842 TTACAACAACATAGACGATGGCTATCAGGCGATCAGACTGTGACGAC-----TCCAGGATG 1898
QY 2214 CACGCCAGACCTGCTATCTACAGAGGGGCCACCGCTGCTCTGCCCAATGTCACGCAGA 2273
Db 1899 CGTGTCTGTTGTGATCGATGGTCCCTCTCCACCAACCCCAACCATCTCTATGCGAA 1958
QY 2274 GTCAACAGCTCCACTTCCATTTGGCTTTCGGTGGAGAGCCAGACTTTTACCACTGTGAA 2333
Db 1959 GGTAAACACCTCATCTTCCATCTTCTGCACTGGAGGAGGCTGCATTTACCGCTGCACA 2018
QY 2334 GATTGTCAACTACACTGTACGTTGGGCCCTTGGGGGCTCAGGAATGCTTCCCTGGTCA 2393
Db 2019 AATCATTAACATACACATCCGCTGTAATCTGTTGGCCCTGCAGAAATGCTTCTTGGTTCT 2078
QY 2394 CTACTATACAGCTCTGAGAGAGACATTTCTATTGGCGGCTGAAACCATTTTACCAAGTA 2453
Db 2079 GTACCTTCAACATCAGAAATCTCATATGTTGGTCAAGGCTAGAAACCAAAACCAATA 2138
QY 2454 CGAGTTTCCGGTACAGTCCACGAGTGGATATGGATGGGCTTCTTGGCTCCGTCGTAGA 2513
Db 2139 CGAATTTGCCGTTGATATGATGATCAGCTTTCAGTCTTGGAGCCCTGTAGTCTA 2198
QY 2514 AGCTTCCACCTGCTGACCGGCTTCAACACCTCTTCTGACCTGGCCCTGAGCCGCCCT 2573
Db 2199 CCAATCTACTCTTCCAGAGCACACAGCAGGCGCCACCACTTGGAGTAAAGTGACATTAAT 2258
QY 2574 GACACCATCCACGTTGGTTTACACTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2633
Db 2259 AGAGATGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2318
QY 2634 GGAGTATCTAATTTCTTACAGCAACACACACACACACACACACACACACACACACACAC 2693
Db 2319 CCGCTATCTATCTTATATGATCTAGGAGGCTGATTTGAGGAGTGGAGGCTT 2378
QY 2694 CACCAAGAGGAAACATCTTCAAGTGCAGAGGTCAATGGCTTACAGAGTGCACCTCGGTA 2753
Db 2379 ACACCGTGAAGGGGCAATAACATGCTTGTAGAAAACCTTGGTAGCAGGAATGTGTA 2438
QY 2754 TTTCTTCAAGATGGAGGCCGACAGAGTGGGGCTGGGGCTTTTCCCGCTTGCAGGA 2813
Db 2439 CATTTGTCAGATATCTGATCCAAATGAGGTGGAGAGGAGCCCTTTTCAAAATCTGTGGA 2498
QY 2814 TGTGATTACTCTGCAAGAGACATTTCTCAGACTC 2846
Db 2499 GCTGGCAGTACTTCCAAAGGAACCTCTGAATC 2531

```

RESULT 9

```

US-10-231-353-1
; Sequence 1, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian

```

```

; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20030023064A1el Human Phosphatases and Polynucleotides En
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231.353
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3210
; TYPE: DNA
; ORGANISM: homo sapiens
; US-10-231-353-1

```

```

Query Match      5.1%; Score 316.6; DB 14; Length 3210;
Best Local Similarity 47.4%; Pred. No. 1.4e-83;
Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6;

QY 300 AGACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCTACTGAGGGCAGCTTATTCCTG 359
Db 279 ATACATCAGTGAAGTGAAGGCAGGCGAGGAGAGCAGTCCGATGAAGGATTTTATCAGTG 338
QY 360 TCTGGCCACACAGCCCGCTAGGAGTGGGCCAGCCAGGTTGCTGTGTGCTCAAGCTTGCAC 419
Db 339 CTTGGCAATGAACAAATATGAGCCATTTCTTAGTCAAAAAGCTCATCTTGCCTTATCAAC 398
QY 420 ACTCGAAGACTTCTCTCTGCAACCCCGAGTCCAGATTTGTGAGGAGAACGGGACGACG 479
Db 399 TATTTCTGCAATTTGAAGTCCAGCCAAATTTCCACTGAGGTCCACGAGGTGGAGTTGCTCG 458
QY 480 CTTTGAATGCAACACCAAGGGCTTCCAGGCCCCCATCATTTCTTGGGAAAGAGACCAAGT 539
Db 459 ATTGTCATGCAAGATTTTCATCCACCCTCTGCACTGATTAACATGAGGATTCATTCGAC 518
QY 540 GACCGTCCCTGAG---GAGCCCGGCTCATCATCTTCCCAAGTGGCTCTTCCAGATCCT 596
Db 519 AACTCTACCTATGACTATGAGACAGGATAACTGCCCTTACCAACAGGAGTATTTGCAGATCTA 578
QY 597 AGATGTCCAGCAGTGTGATGATGAGGCTCTTACCGCTGCTGCGCCACCAATTCAGCCCGCA 656
Db 579 TGATGTGAGCCAAAGGATTTCTGGAATTTATGTTGTTGTTGCTGCTGCTGAGGCCACCG 638
QY 657 ACGATTGAGCCAGAGCCCTCGCTCAGTGGGCCCTCAGAGGCTTTTGGAGGCTTACGAG 716
Db 639 ACGTAAAGTATGAGGCGCTCGCTAACTGTGATTTCCAG-----CTAAGGAGTCAAAATC 692
QY 717 GGGGCGAGATGTGCTCATTTGTGCGAGCCCGCAGAGAACACCAAGTGTGTGACAGAA 776
Db 693 CTTCCACACACACCAATTTATAGCAGGTCCACAGAAATTAACAAATCTCTTCTCATGAC 752
QY 777 TGTAGTGTGAGTGTGCTGCTGCTGACCCCAACCCCTTTTGTGCTGCTGCTGCTGCTGCT 836
Db 753 TGTAGTTTGGATGATGAGCCACAGGAATCCCAACCAATCATTTCTTGGAGCGGCT 812
QY 837 GGTGGAAGGCTTAT---CTCCACGAGTGTATCGTTCTGCGGCCGAGCAATCTACTCAT 893
Db 813 TGATCACAATTCATGATGCTTTTAACTCTGCGGTACTTGGAAATGTTATCTCATGAT 872
QY 894 GCGGAGCGGCGAGCTCGGCACTGTGAGTGTATGCTGTGCGAGCCCAACAGCCCTCAC 953
Db 873 ATCTGATGTGAGGCTACAAATGCTGGAGTATATGTTTGTGGGCCCACTACCCCTGGCAC 932
QY 954 GCGTGAATTCGCCACTGCGGCTGCTGAGTCCGAGTGTGCTGCGCCCGCAGCATCTCGCA 1013
Db 933 ACGCACTTTTACAGTTCTATGGAATTTTAACTGTTATAGTCTCTCTCTCTCTCTCTCTCT 992
QY 1014 GGCAACCCGAGGCGCTCTCGGGGAGCGGGGCGAGCACCGCGCTTCTGTTGCTGCGGGGCTC 1073
Db 993 ATGGCCAGAAAGTTTAAACAGGCTCGAGCTGGCACTGCTCGATTTGTGTGTGAGGAGCA 1052
QY 1074 CGGGGAGCCACGGCCCGCTGCACTGCTGTCAGCAGCGGATCCCGTTGCGACCAATG 1133

```

Db 1053 AGGAATCCCTCTCCCAAGATGTCATGGTTGAAAAATGGAAGGAGATACATTCGAATGG 1112
Qy 1134 GCCTGCAAGGTGACGGGGTGGCGGACGCTTGGTCATCACTCAGATCGGCTGCAGGA 1193
Db 1113 TAGAATTAATGATACACAGT-----AAATGGTAAATTAACCAAGATTAATTCCTGAAG 1166
Qy 1194 CGCTGCTACTACCACTGGTGTAGCAGAAACAGCGCGGAACATGCTGTGCGCTGCGCC 1253
Db 1167 TGATGCTATTTATCACTGATGCTGAGAAATAGCCAAAGATCTATTTATCTAGAGCCAG 1226
Qy 1254 CTTGGCGGTAGTGGTGGGAGGGGCTGCCAGCGGCCGACATCGGGTCAACAGCCAGCC 1313
Db 1227 ACTGACTGTAGTGTATGAGAAGACAGACCCAGTGTCTCCCTATAATGTATCATGCTGAAC 1286
Qy 1314 GCTGAGCAGCTCTCTGTGCTGGTGGCTGGGAGCGCTGAGTTGCAAGCGAGCAAT 1373
Db 1287 CATGTCAAGCTCAGCCATTTCTTTAGCTGGGAGGCCACTTTATTAATTCAGACAAAGT 1346
Qy 1374 CATTTGCTTCTCTCTTCACTACCAAAAGGAGGGGAGTGGACAAATGTGGAGTACCAATT 1433
Db 1347 CATTTGCTTCTCTGTAACCTACATGAAGAGGAGGTTTAAATAATGAAGAGTATCAAGT 1406
Qy 1434 TGCAGTAAACAAATGACACACAGAGCTGCAGGTTTGGGACCTGGAAACCCAAACAGGATTA 1493
Db 1407 AGTCATCGGAATGACACAACTCATTAATTAATGATGACTTAGAGCTTGCACGCAATTA 1466
Qy 1494 TGAAGTTCTAGCTGGTCTTACTCCAGCTGGGGGCGAGCCGAAACCTCCAGGCCAGCCCT 1553
Db 1467 TACTTTCTATCTGTAGTATATATGCAATGGGAGCCAGCAGATGTCTGACCATGTGAC 1526
Qy 1554 GGTGCATACATGGAGCATGTCCCGAGCGAGCACCCCGAGCTTACCTTTGTCAGCCCCAA 1613
Db 1527 ACAGAACTCTAGAGGATGTTCCCTCTGAGACCTCTCGAAATTTAGTTTGACAAAGT 1586
Qy 1614 CCCTCGGACATCAGGGTGGCTGCTGCCCTGCTCCCTGAGCTGAGCAATGGACAGT 1673
Db 1587 TCCCACTGATTAATCTCATCTCTGCTGCTGCAATCCAGCCAAATATCGGGCGGGCAGT 1646
Qy 1674 GCTGAAGTAAAGATAGATGACGTTTGGGAGGAGGATGAGTTTCTCCACCGAGT 1733
Db 1647 GGTGCTGATCGCTGCTCTTCCGCTTAAGTACTGAGAAATCAATCAAGTTCTGAGCT 1706
Qy 1734 GCCTGAAATAGACACAACTTACGTAACTCACTTCAGCCCAACAAAGTGTACCGAGT 1793
Db 1707 CCGGGGACCCAGCATGATGATCTTTTGAAGGCTGAAACCTGACAGTGTCTACCTGCT 1766
Qy 1794 CCGAATTCAGTGGACATCGGCTGCTATGAGTCCCTTCTCAGTGGATGACAGCAG 1853
Db 1767 TCGGATTTACTGCTGCCACCAAGAGTGGGCTGGGAGAGTATCAGTATGGAATTCACATAG 1826
Qy 1854 GACACCTGGTGTGCACAAACAGAGCCATGTTCCCTTTGCCCTGCAAGAAATGAAGTGAG 1913
Db 1827 GAGCCCAAGACTACAAGCGTGAAGCCCTTAAGTCTCAGAGTTGATTTGAGGCTCT 1886
Qy 1914 GGCAAGATGAGTCCCTGCTGTGTATGCGACGCGCCCTTCAACCCCAACCAAGATCTC 1973
Db 1887 GAACTGTACCACATTTCTGTGAGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCA 1946
Qy 1974 TGATTAACAACTCTACTGGGAGAGGTGGAAACAGAGGAGGAGGAGATGTTGACCGCC 2033
Db 1947 GGGCTACAAGCTGTACTACAAGGAAGAAAGGCGACGAGAA----- 1988
Qy 2034 CCAGGGGGTGTGGAGATCAAGCTTGGGAGCTGGGCGCGCTGAGGAGAAAGT 2093
Db 1989 -----TGGGCCCATTTTCTTGGATACCAAGGA 2015
Qy 2094 GAAGCAGTATGAATGACCCAGTTAGTCTCTGGCAGCGCTGACGAGTGAAGTCTGTAGC 2153
Db 2016 CCTACTCTACTCTCAGTGGCTTAGACCCCAAGAAATATCATGTGAGACTCTCTGGC 2075
Qy 2154 TTTCAACAAACACAGAGCAGGCTACGCTGTGTGTGGAAGGGCAAGACGGAGAGCGCC 2213
Db 2076 TTACAACAAATAGACATGGCTATCAGGAGATCAGACTGTGAGCAC---TCCAGGATG 2132

Qy 2214 CACGCCAGACCTGCCTATCCAGAGGGGGCCACCGCTGCTCTCTGCCCATTGTCACGACGA 2273
Db 2133 CGTGTCTGTTCGTGATGCAATGGTCCCTCTCCACCAACCCCAACCATCTCTATGCGAA 2192
Qy 2274 GTCAAAAGCTCCACTTCCATTTGGCTTGGTGGAGGAGGAGCAGACTTTTACACATGTCAA 2333
Db 2193 GGTAAACACCTCATCTTCCATCTTCTGCACTGGAGGAGGCTGCAATTCACCGCTGCACA 2252
Qy 2334 GATTGTCAACTACACGTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCCTGGTCA 2393
Db 2253 AATCAATTAACATACCACTCCGCTGTAATCTTGGTGGCTTGAGAAATGCTTCTTGGTCT 2312
Qy 2394 CTACTATACCACTCTGGAGAGACATTTCTCAATTTGGCGGCTGAAACCAATTTACCAAGTA 2453
Db 2313 GTACCTTCAAAACATCAGAAACCTCACATGTTGGTTCAAGGTCTAGAAACCAACCAATA 2372
Qy 2454 CGAGTTTGGGTACAGTCCCAAGGAGTGTATGATGGGCGCTTTGGCTCGCTGTGATA 2513
Db 2373 CGAATTTGCGGTTGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 2432
Qy 2514 AGCTTCCACCTGCTGACCGCTTCAACACCTCTTCTGACCTGCGCTGAGCCCT 2573
Db 2433 CAATTTCTACTCTTCCAGAGGACGAGCGCCACAGTTGAGTAAAGTGACATTAAT 2492
Qy 2574 GACCAATCCACCTGCTGCTTACACTGGTGTCCCGGAGGCCCAATGGTGAATGT 2633
Db 2493 AGAGGATGACACTGCGCTGTTCTTGGAAACCCCTGATGGCCAGAAACAGTTGTGAC 2552
Qy 2634 GGAGTATCTAATTTCTTACAGCAACACACACCGCCCGAAACCAAGTGGACACTGCT 2693
Db 2553 CGCTATATCTATCTATATGATCTAGGAAGCTTGAATTCAGGAGAGTGGCAGTCTT 2612
Qy 2694 CACCAAGAGGAGAAACATCTTCAGTGCAGAGTCCATGCGCTAGAGAGTGCACCTCGTA 2753
Db 2613 ACACGTAAGGGGCAATAACCATGCTTTGTAGAAACTTGGTAGCAAGAAATGTGTA 2672
Qy 2754 TTTCTTCAAGATGGAGCGCCGACAGAGTGGGGCTTGGGCGCTTTTCCCGCTTGCAGGA 2813
Db 2673 CATTTCAAGATATCTGATCCATGAGTGGGAGAGGAGCCCTTTTCAAAATTCGTGGA 2732
Qy 2814 TGTGATTAATCTGCAAGAGACATTTCTCAGACTC 2846
Db 2733 GCTGCGAGTACTTCCAAAGGAAACCTCTGAATC 2765

RESULT 10

US-10-231-353-17
; Sequence 17, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20030023064A1el Human Phosphatases and Polynucleotides E
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231,353
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3219
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-231-353-17

Query Match

5.1%; Score 316.6; DB 14; Length 3219;

Db 1717 ACAGAAATCTAGAGGATGTTCCCTGAGACCTCTCTGAAATTTAGTTTGACAAGTCGAAG 1776
Qy 1614 CCCCTCGACATCATGGGTGGCATGCTGCCCTCGCCCTCCAGCTCGAGCAATGACAGGT 1673
Db 1777 TCCCACTGATATCTCATCTCTGACTGCCAATCCAGCCAAATATCGCGGGGCCAAGT 1836
Qy 1674 GCTGAAGTACAGATAGAGTACGGTTTGGGAGAGAGATACAGGTTTCTCACCGAGGT 1733
Db 1837 GGTGCTGATCGCTTGTCTTCCCGCTAAGTACTGAGAAATTCAAATTCGAAGTTCTGGAGCT 1896
Qy 1734 GCCTGGAATGAGACACAACTTACGTTAACTCACTTCAGCCCAACAAAGGTGACCGAGT 1793
Db 1897 CCCGGGACCAAGCATGATGATCTTTTGGAGGCTGAAACCTGACAGTGTCTACTCGT 1956
Qy 1794 CCGGATTTTCAGTCTGCGCTGCGCTATGGAGTCCCTTCTCAGTGGATGCGACACAG 1853
Db 1957 TCGGATTAAGTCTGCCACAGAGTGGGCTGGAGAGTCACTAGTATGACTTCACATAG 2016
Qy 1854 GACACCTGGTGGCAACACAGAGCCATGTTCCCTTTTCCCTTCCCTGCAAGATTAAGAGTGAAG 1913
Db 2017 GACGCCCAAGCTACAGCGCTGAAGCCCTAAGTCTCCAGAGTTGCAATTTGGAGCCTCT 2076
Qy 1914 GGCAAGATGAGTCCCTGCTGCTGCTGATGGCAGCGGCCCTCACCCACCCAGATCTC 1973
Db 2077 GAACTGTACCACTATCTGTGAGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCA 2136
Qy 1974 TGGATACAACTCTACTCGGAGAGGTGGACAGAGGAGGAGGAGGAGGAGTGTGACCGGCC 2033
Db 2137 GGGCTCAAGCTGTACTACAGGAAGAGGGCAGCAGAGAA----- 2178
Qy 2034 CCCAGGGGTGCTGGAGATCAAGCTTGGGACGTGCGGCCCGTGGCGCTGGAAGAGAAAGT 2093
Db 2179 -----TGGGCCCATTTTCTTGGATACCAAGGA 2205
Qy 2094 GAAGCAGATAGAACTGACCCAGTTAGTCCCTGGCAGGCGGTACGAGGTGAAGCTGTAGC 2153
Db 2206 CCTACTCTATATCTCTCAGTGGCTTAGACCCCAAGAAATAATCATGTGAGACTCTCTGC 2265
Qy 2154 TTTCAACAAACACAGGAGGCTGCTGCTGTGTGGAGGCGAGACGGAGAGGCGCC 2213
Db 2266 TTACAACAAACATAGACAGATGGCTATCAGGCGATCAGACTGTGACAC-----TCCAGGATG 2322
Qy 2214 CACGCCAGACCTGCTATCTCAGAGGGGCGCACCGCTGCTCTGCCCCATGTCCACGACGA 2273
Db 2323 CGTGTCTGTGATGATGATGCTCTCTCCACCAACCCCACTCTCTATGCGAA 2382
Qy 2274 GTCAACAGCTCCACTTCCATTTGGCTTTCGTTGGAGAGGAGGAGCTTTTACACTGTCAA 2333
Db 2383 GGTAAACACCTCATCTTCCATCTTCTGCACTGGAGGAGGCTGCAATTCACCGCTGCACA 2442
Qy 2334 GATTGTCACTACACTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCCTGGTCAAC 2393
Db 2443 AATCATTAACATACACATTCGCTGTAATCTGTTGGCCCTGACGAATGCTTCTTTGGTTCT 2502
Qy 2394 CTACTATACCACTCTGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTA 2453
Db 2503 GTACCTTCAACATCAGAACTCAATGTTGGTTCAAGTCTAGAACCAACCAACCAATA 2562
Qy 2454 CGAGTTTCGGTACAGTCCACGAGTGGATATGGATGGGCCCTTTTGGCTCGTGTAGA 2513
Db 2563 CGAATTTCCGTTCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2622
Qy 2514 AGCTTCACCTGCTGACCGGCTTCAACACCTCTCTGACCTGCGGCTGAGCCCTT 2573
Db 2623 CCAATTCATCTCTTCAGAGACACACAGGCGGCCACCACTGTGAGTAAAGTGACATTAAT 2682
Qy 2574 GACACCATCCAGTTCGTTTACACTGTGTGTCCTCCCGCAGGAGCCCAATGTTGATGTT 2633
Db 2683 AGAGATGACATGCTGCTGTTTCTTGGAAACCCCTCTGATGGCCCAAGAAACAGTTGTGAC 2742
Qy 2634 GGAGTATCTAATTTCTACAGCAACACACACACCCAGCCGCAACCAACCACTGTGACACTGCT 2693
Db 2743 CCGCTATACTATCTATATGATCTATAGGAGGCTGGATTCAGAGAGATGCGAGGCTCTT 2802

Qy 2694 CACACAGAGGAAACATCTTCAAGTCAGAGGTCATATGGCTAGAGAGTGAACACTCGGTA 2753
Db 2803 ACACCGTGAAGGGCAATAACCAATGCTTTGCTAGAAAACCTTGTAGCAGGAATGTGTA 2862
Qy 2754 TTTCTTCAAGATGGAGCCCGCACAGAGTGGGCTTGGGCCCTTTTCCCGCTTGCAGGA 2813
Db 2863 CATTTGTCAGATATCTCATCCATGAGGTGGGAGAGGACCCCTTTTCAAAATTCGTGGA 2922
Qy 2814 TGTGATTAATCTGCAAGAGACATTTCTCAGACTC 2846
Db 2923 GCTGGCAGTACTTCCAAAGGNAACCTCTGAAATC 2955

RESULT 12

US-10-231-353-7
; Sequence 7, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20030023064A1el Human Phosphatases and Polynucleotides En
; CURRENT APPLICATION NUMBER: US/10/231,353
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3453
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-231-353-7

Query Match 5.1%; Score 315.4; DB 14; Length 3453;
Best Local Similarity 47.4%; Pred. No. 3.4e-83;
Matches 1210; Conservative 0; Mismatches 1271; Indels 72; Gaps 6;

Qy 300 AGACAGGATGATGAGGAGCTCTTAGATCTGGAAGTCACTGAGGCGAGCTATTCCTG 359
Db 279 ATACATCAGTGGTGAAGGACGCGGAGGAGAGCAGTCCGATGAAGATTTTATCAGT 338
Qy 360 TCTGGCCACAGCCCGCTAGGAGTGGTGGCCAGCCAGGTTCTGTGTCAAGCTTGCAC 419
Db 339 CTGGCAATGAAACAATATGAGGCCATTTCTTAGTCAAAAGCTCATCTTGCCTTATCAAC 398
Qy 420 ACTCGAAGACTTCTCTCTGCAACCCGAGTCCAGATTTGTGGAGAGAAACGAGCAGC 479
Db 399 TATTTCTGCAATTTGAAGTCCAGCCAAATTTCCACTGAGGTCCAGGAGGTGAGTGTCTG 458
Qy 480 CTTTGAATGCCACCAAGGCGCTTCCAGCCCCCATCATTAATTTGGGAAAGGACCGGT 539
Db 459 ATTTGCAATGCAAGATTTTCATCCCACTCTCTGAGTCAATAATGAGGATTCATTCGAC 518
Qy 540 GACCGTCTCTGAG--GAGCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCCT 596
Db 519 AACTCTATGATGATGAGCAGGATTAATGAGGCTCAACACAGGAGTATTCAGATCTA 578
Qy 597 AGATGTCCAGGACAGTGTGAGGCTCTCTACCGTGGTGGCCACCAATTCAGGCCGCCA 656
Db 579 TGATGTGAGCCCAAGGATTTCTGGAAATTTATGTTGTTATTTGCTGCACTGTAGCCCA 638
Qy 657 ACATTTGAGCCAGAGGCTCTGCTCACTGTGGCCCTCTGAGGGGTCTTTTGGAGGTCTAC 716
Db 639 ACGTAAAGATATGGAGGCTCTGCTAACTGTGATTTCCAG-----CTAAGGAGTCAAAATC 692
Qy 717 GGGGAGGATGTGTGATTTGTGGAGCCGCCAGAGAACACACCGGTAGTGTCTGGACAGAA 776

Db	693	CTTCCACACACCAACAAATATAGCAGGTCCACAGAAATAAACAAATCTCTCATCAGAC	752	1767	TCGGATTACTGCTGCCACACAGAGTGGGGCTGGGAGAGTCAATCAGTATAGACTTTCATATAG	1826
Qy	777	TGTAGTATGAGTGGTGGCTCTGCTGACCCACCCCTTTTGTGCTGGGTCCGACA	836	1854	GAACCTGTGTGTCACAAACAGAGCATGTTCCCTTTTGCCTTGCAGAAATTCGAAGTGGAG	1913
Db	753	TGTAGTTTGGAAATGCAATGGCCACAGAAATCCCAACCAATCAATTTCTTGGAGCGCCT	812	1827	GAAGCCCAAGGCTACAGAGGTGAAGCCCTTAAGTCTCCAGAGTTGCATTTGGAGCCTCT	1886
Qy	837	GGATGGAAGCCTAT---CTCCACGGATGTATCGTTCTGGGCGGACCAATCTACTCAT	893	1914	GGAAAGATGGAGTCCCTGGTGTGTATGGCAGCGGCCCTCCACCCACCCAGATCTC	1973
Db	813	TGATCACAAATCCATTCATGCTCTTAATCTCGGGTACTTGGAAATGGTAATCTCATGAT	872	1887	GAACCTGTACCAACATTTCTGTGAGGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCA	1946
Qy	894	CGCAGCGCGAGCTCGGCACTCTGGAGTCTATGTCTGCGAGCCAAAGCCCTCAC	953	1974	TGGATACAAATCTACTGGGAGAGTGGGAAACAGAGGAGGAGGAGATGGTGACCGCCC	2033
Db	873	ATCTGATGTGAGGCTACAAATGCTGGAGTATATGTTGTGCGGCCCACTACCCCTGCGAC	932	1947	GGGCTACAGCTGTACTACAGGAAGAGGGCAGCAGAGAA-----	1988
Qy	954	GCCTGACTTCGCCACTCGGCTGCTGAGCTCGGAGTCTTGTGCCCCCAGCCATCTCGCA	1013	2034	CCAGGGGTCTGGAGATCAAGCTTGGGACGTCCGGCCCTGGGCTGGAAGAGAAGT	2093
Db	933	ACGCACTTTTACAGTTGCTATGGCACTTTAACTGTATTAAGTCTCTCTCATTTGTTGA	992	1989	-----TGGGCCCATTTTCTTGGATACCAAGCA	2015
Qy	1014	GGCACCGAGCGCTCTCGGGGACGCGGGCCAGCACCGCGCTTCTGTGCGGGCGTC	1073	2094	GAAGCATATGAATGACCCAGTTAGTCCCTGGAGCGGTATAGAGGTGAAAGCTGTGTAGC	2153
Db	993	ATGSCCAAGAAAGTTTAAACAAAGGCTCGAGCTGGCAGCTGCTCGATTTGTGTGTGAGCAGA	1052	2016	CCTACTCTATCTCTCAGTGGCTTAGACCCCAAGAAATAATATCATGTGAGACTCCTGCGC	2075
Qy	1074	CGGGAGCCAGCGCCCGCTGCACTGGCTGCGAGCGGATCCGTTGGCAGCCCAATGG	1133	2154	TTTCAACAAACAGGAGCGGCTACGCTCTGTGTGGAAGGCAAGACGGAGAGGCGCC	2213
Db	1053	AGGAATCCCTCTCCCAAGATGTATGTTTGAATAATGGAAGGAAGATACATTCGAATGG	1112	2076	TTACAAACAATAGACGATGGCTATCAGGCAGATCAGACTGTGAGC-----TCCAGGATG	2132
Qy	1134	GCAGTCAAGTGCAGGCGGCTGGCGGAGCTTGGTCACTCAGATCGGCTCGAGGA	1193	2214	CACGCCAGACCTGCGCTATCCAGAGGGGCCACCGCTGCTCTTGCCTCATGTGCCACAGCA	2273
Db	1113	TAGAAITAAATGTACAAACAGT-----AAATGGTAAITAAACAGATTAATCTCTGAAGA	1166	2133	CGTGTCTGTGTGTGATCGCATGGTCCCTCTCCACCAACACCCCACTCTCTATGCGAA	2192
Qy	1194	CGCTGGCTACTACAGTGCCTAGCAGAAACAGCGCGGAACCTGCTGTCGCGCTGGCC	1253	2274	GTCAAAACAGCTCATCTTCCATTTTGGCTTCGGTGGAAAGAACGACACTTTTACACTGTGCA	2333
Db	1167	TGATGCTATTTATCAGTGCATGGCTGAGATAGCCAGGATCTATTTATCTAGAGCCAG	1226	2193	GGCTAACACTCATCTTCCATCTTCTGCACTGGAGAGGCTTGCATTCACCGCTGCACA	2252
Qy	1254	CCTGGCGGTAGTGGTGCAGAGGCTGCCAGCGGCCCACTGCGGTCAAGCCAGCC	1313	2334	GAATGTCAACTACACTGTAGCTTGGCCCTTGGGGCTCAGGAATGCTTCCCTGCTCAC	2393
Db	1227	ACTGACTGATGATGTGAGAAAGACAGACCCAGTGTCTCTTAATGTATCATGTGAAC	1286	2253	AATCATTAATACACCATCCGCTGTAATCTTGTGGCTTCGCAAGATGCTTCTTGTGTTCT	2312
Qy	1314	GCTGAGCAGCTCTCTGTGTGTGGCTGGAGCGGCTGAGTTGCAGACCGCAAT	1373	2394	CTACTATACAGCTCTGGAGAAAGACATCTCTCATTTGGCGCTGAAACCAATTTACCAAGTA	2453
Db	1287	CATGTCAAGTTCAGCCATCTTTTAGCTGGAGAGGCCACTTTATTAATTCAGACAAAGT	1346	2313	GTACCTTCAAAATCAGAAACTCACAATGTTGGTTCAGGTCTAGAACCAACCAATA	2372
Qy	1374	CATTGGCTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGCAATGTGGATACCAAGTT	1433	2454	CGAGTTTGGGTACAGTCCCAAGAGTGATATGGATGGGCCCTTTGGCTCGGTGTAGA	2513
Db	1347	CATTGGCTATCTGTACACTACATGAAGCAGAGGTTTAAATTAATGAAGATCAAGT	1406	2373	CGAAATTTGGCGTTGATTAATATGTGATCAGCTTTCAGTCTTGGAGCCCTGTAGTCTA	2432
Qy	1434	TGCAGTAAACAAATGACACCAAGAGCTGCGAGTTTCGGGACCTGGAAACCAACCGGATTA	1493	2514	ACGCTCAACCTGCTGACCGGCTTCAACACTCTTCTGACCTGCGCTGAGCCCCCT	2573
Db	1407	AGTCATCGGAAATGACACAACTCAATTAATTAATGATGACTTAGAGCTGCGAGCAATTA	1466	2433	CCATTCCTCTTCCAGAAAGCACAGAGGCCACCAAGTTGGAGTAAAGTGCATTAAT	2492
Qy	1494	TGAGTTCTAGTGGTGGCTACTCCAGCTGGGGCCAGCGCAACCTCCAGCGCCCT	1553	2574	GACACCATCCCGTTCCGTTTACACTGTGTGTCCTCCCAAGGCCCAATGGTGCAGATTGT	2633
Db	1467	TACTTTCTACATTTGATATATATGCAATGGGAGCCAGCCAGATGTCTGACCATGTGAC	1526	2493	AGAGGATGACACTGCGCTGTTTCTTGGAAACCCCTGTATGGCCAGAACAGTTGTGAC	2552
Qy	1554	GGTCATACACTGACCATGTCCCGAGCGAGCACCCAGCTTACCTTGTCCAGCCCAAA	1613	2634	GGAGTATCTAAATTTCTTACAGCAACAAACACACCCAGCCGGAACACAGTGTGCT	2693
Db	1527	ACAGAACTCTAGAGATGTTCCTCTGAGACCTCTCGAAATAGTTTGAAGTCTGAAG	1586	2553	CGCTATATCTATCTTATATGCACTTAGGAAGGCTGGAATTCAGAGAGAGTGGCAGTCTT	2612
Qy	1614	CCCCCGACATCAGGGTGGCAGTGGCTGCCCTGCCCTCCAGCTGAGCAATGACAGGT	1673	2694	CACCACAGAGGAAACATCTTCACTGTCAGAGGTTCATGGCCTTACAGAGTGCACACTCGGTA	2753
Db	1587	TCCCACTGATATCTCACTCTCGCTGCCAAATCCAGGCCAAATATCGCGGGGCCAAGT	1646	2613	ACACCGTGAAGGGGCAATAACCAATGGCTTTGTAGAAACCTTGGTAGCAGGAATGTGA	2672
Qy	1674	GCTGAAGTACAGATAGGTACGGTTTGGGGAAGGAAGATCAGGTTTTTCTCCAGAGGT	1733	2754	TTTCTTCAAGATGGGAGCCGACAGAGTGGGGCTGGGCCCTTTTCCGCTTTCAGGGA	2813
Db	1647	GGTGTCTATGCTGTGTTCTTCGCTTAAGTACTGAGAAITCAATCCAAAGTTCTGGAGCT	1706	2673	CATTGTCAAGATATCTGCAATCCATGAGGTGGGAAGGAGGCCCTTTTCAATTTCTGTGGA	2732
Qy	1734	GCCTGGAATAGAGACAACTTACGTTAAATCACTTTCAGCCAAACAAAGTGTACCGAGT	1793	2814	TGTGATTACTCTGCAAGAGACATTTCTCAGACTC	2846
Db	1707	CCCGGAGCCACGCACTAGTACCTTTTGGAGGCTGAAACCTGACAGTGTCTACCTGGT	1766	2733	GCTGGCAGTACTTCCAAAGGAACCTCTGAATC	2765
Qy	1794	CCGATTTTCACTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGGATGCAGACAG	1853			

```
; Sequence 19, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(288)
US-09-754-997A-19

Query Match
Best Local Similarity 4.7%; Score 288; DB 10; Length 288;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1888 TTTCGCCCCGTCAGATTGAAGTGGAGGCAAGATGGAGTCCCTGGTGTCTATGGCAG 1947
Db 1 TTTCGCCCCGTCAGATTGAAGTGGAGGCAAGATGGAGTCCCTGGTGTCTATGGCAG 60

QY 1948 CGCCCCCTCCACCCACCCAGATCTCTGGATCAAACTCTACTGGGGAGAGGTGGGAACA 2007
Db 61 CGCCCCCTCCACCCACCCAGATCTCTGGATCAAACTCTACTGGGGAGAGGTGGGAACA 120

QY 2008 GAGGAGGAGGAGATGGTGACCGCCCCCGGGGTCTGGAGATCAAGCTTTGGACGTC 2067
Db 121 GAGGAGGAGGAGATGGTGACCGCCCCCGGGGTCTGGAGATCAAGCTTTGGACGTC 180

QY 2068 GGGCCGTCGCGCTGAAGAGAGAGTGAAGCAGTATGAATGACCCAGTTAGTCCCTGGC 2127
Db 181 GGGCCGTCGCGCTGAAGAGAGAGTGAAGCAGTATGAATGACCCAGTTAGTCCCTGGC 240

QY 2128 AGCGCGTACGAGTGAAGTCTGTAGCTTTCAACAAACACGAGGAGCGC 2175
Db 241 AGCGCGTACGAGTGAAGTCTGTAGCTTTCAACAAACACGAGGAGCGC 288

RESULT 14
US-09-754-997A-23
; Sequence 23, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 252
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(252)
```

```
US-09-754-997A-23

Query Match
Best Local Similarity 4.1%; Score 252; DB 10; Length 252;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2542 ACACCTCTCTTGACCTGGCTGAGCCCTCGACACCATCCACCGTTCCGTTACACTGG 2601
Db 1 ACACCTCTCTTGACCTGGCTGAGCCCTCGACACCATCCACCGTTCCGTTACACTGG 60

QY 2602 TGTCCCCCAGCGAGCCCAATGGTGAGATTCTGGAGTATCTAATTTCTTACAGCAACAC 2661
Db 61 TGTCCCCCAGCGAGCCCAATGGTGAGATTCTGGAGTATCTAATTTCTTACAGCAACAC 120

QY 2662 CACACCCAGCCCGAACCACTGCTCACCACAGAGGAAACATCTTCAAGTCA 2721
Db 121 CACACCCAGCCCGAACCACTGCTCACCACAGAGGAAACATCTTCAAGTCA 180

QY 2722 GAGGTCCATGGCTTAGAGAGTGACACTCGGTATTTCTTCAAGTGGAGCCCGCACAG 2781
Db 181 GAGGTCCATGGCTTAGAGAGTGACACTCGGTATTTCTTCAAGTGGAGCCCGCACAG 240

QY 2782 GTGGGGCTGGG 2793
Db 241 GTGGGGCTGGG 252

RESULT 15
US-09-754-997A-15
; Sequence 15, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(249)
US-09-754-997A-15

Query Match
Best Local Similarity 4.0%; Score 249; DB 10; Length 249;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1285 AGCGCCCGACTCGGTGACAGCCAGCCGCGCTGAGCAGCTCTCTGTGCTGGCTGG 1344
Db 1 AGCGCCCGACTCGGTGACAGCCAGCCGCGCTGAGCAGCTCTCTGTGCTGGCTGG 60

QY 1345 GAGCGGCTGAGTTGCAAGGAGCAATCAATGGCTTCTCTTCACTACCAAGGCA 1404
Db 61 GAGCGGCTGAGTTGCAAGGAGCAATCAATGGCTTCTCTTCACTACCAAGGCA 120

QY 1405 AGGGAGTGGCAATGTGGAGTACCAAGTTTGAGTAAACATGACACCAAGAGCTGAG 1464
Db 121 AGGGAGTGGCAATGTGGAGTACCAAGTTTGAGTAAACATGACACCAAGAGCTGAG 180

QY 1465 GTTCGGGACCTGGAAACCCAAACACCGATTAAGATTCTACGTGGTGGCTTCTCCAGCTG 1524
Db 181 GTTCGGGACCTGGAAACCCAAACACCGATTAAGATTCTACGTGGTGGCTTCTCCAGCTG 240
```

Qy 1525 GGGGCCAGC 1533
| | | | |
Db 241 GGGGCCAGC 249

Search completed: October 15, 2003, 12:28:06
Job time : 1679.34 secs

Qy	597	AGATGTCAGACAGTGAATGAGGCTCCTACCGCTGGTGGCCACCAATTCAGCCGCCA	656	1413	GGTGTGTATCGCTGTGTCTTCCGGCTAAGTACTAGAAATTCATTCAGTTCGTGAGCT	1472	
Db	345	TGATGTGAGCCAAAGGGAATCTGGAAATTAATGTTGTAATTCGTGCTGACCTAGCCACCG	404	Qy	1734	GCCTGAAATAGACACAACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGT	1793
Qy	657	ACGATTCAGCCAGAGGCTCGCTCACTGTGGCCCTCAGAGGCTCTTGGAGGCTACCG	716	Db	1473	CCGGGGACCAAGCATGAGTACTTCTTGGAGGCTGAAACCTGACAGTGTCTACCTGCT	1532
Db	405	ACGTAAGATATGAGGCTCGCTAACTGTGATTCAG-----CTAAGGAGTCAAAATC	458	Qy	1794	CCGATTTCACTGAGCTGGGCTATGGAGTCCCTCTCTCAGTGTGATGACGACAG	1853
Qy	717	GGGCGAGATGTGCTCATTTGGCAGCCCAAGAAACACACGCTAGTGTCTGACAGAA	776	Db	1533	TGGATTAATCTGCTGCCACAGAGTGGGCTGGGAGAGTCAATGATGACATTCACATAG	1592
Db	459	CTTCCACACACCAATATAGCAGGTCCACAGAAATTAACACATCTCTCATCAGAC	518	Qy	1854	GCACCTGTGTGACACACAGGCAATGTTCCCTTTGGCCCTTCAGAAATTCAGAGTGA	1913
Qy	777	TGTAGTGAATGAGTGGCTCTGCTGACCCCAACCCCTTTTGTGTGCTGGGTCCGACA	836	Db	1593	GACGCCAAAGCTACAAGCGTGAAGCCCTTAAGTCTCCAGAGTTGCTATTTGGAGCTCT	1652
Db	519	TGTAGTTTGGAAATGCAATGACACAGGAATTCACCAACCAATCATTTCTTGGAGCCCT	578	Qy	1914	GGCAAGATGAGTCCCTGTGTGTCATGCGAGCCGCCCTCCACCCCAACAGATCTC	1973
Qy	837	GGATGGAAGCCTAT---CTCCAGGATGTATCGTTCTGGCCCGGACCAATCTACTCAT	893	Db	1653	GAACTGTACCAACATTTCTGTGAGGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCA	1712
Db	579	TGATCACAAAATCCATGATGTCTTTAATCTCGGGTACTTGGAAATGGTAAATCTCATGAT	638	Qy	1974	TGGATACAAATCTCTACTGGGGAGAGTGGGAAACAGAGGAGGAGGAGATGGTGAACGCC	2033
Qy	894	CGCCAGGCGCAGCTCGGCACTCTGGAGTCTATGTCTGCGAGCCCAACAGCCCTCAC	953	Db	1713	GGGCTACAAGCTGTACTACAGGAAGAGGGCAGCAGAGAA-----	1754
Db	639	ATCTGATGTAGGCTCAACATGCTGGAGTATATGTTGTGCGGCCACTACCCCTGGCAC	698	Qy	2034	CCAGGGGTCTGGAGATCAAGCTTGGGACGTGGGCGCGCTGGGCTGAAGAGAAAGT	2093
Qy	954	CGGTGACTTCGCCACTCGGCTGCTGAGCTCCGAGTCTGCTGCCCGACCAATCTCGCA	1013	Db	1755	-----TGGGCCCATTTTCTTGGATACCAAGGA	1781
Db	699	ACGCAACTTTACAGTGTATATGGCAACTTTAACTGTATTAGTCTCTCTTCAITTTGTGA	758	Qy	2094	GAAAGATATGAATCAGCCAGTGTAGTCTCTGCGAGCGCTACGAGTGAAGTCCGTAGC	2153
Qy	1014	GGCACCGGAGCGCTCTCGGAGCGCGGCGCAGCACCGCGCTCTGCTGTGCGGGGCTC	1073	Db	1782	CCTACTCTATCTCTGAGTGTAGACCCAGAGAAATATCATGTGAGATCTCTGGC	1841
Db	759	ATGGCCAGAAAGTTTAAACAGGCTCGAGCTCGCACTGCTCGATTTGTGTCTAGGCGA	818	Qy	2154	TTTCAAACAAACAGCAGGCGCTACGCTGTGTGTGGAAGGGCAAGACGGAGAGGGGCC	2213
Qy	1074	CGGGAGCCAGCGCGCGCTGCACTGCTGCGACGAGGATCCGTTGGACCCCAATGG	1133	Db	1842	TTACAACAACATAGACATGGCTATCAGGCAGATCAGATGTCTCAGCAC---TCCAGGATG	1898
Db	819	AGGAATCCCTCTCCAGATGTATGTTGAAATGGAAGAGATACATTCGAATGG	878	Qy	2214	CAGCCAGACCTGCTTATCAGAGGGGGCCACCGCTGCTCTGCGCCATGTCCACGAGA	2273
Qy	1134	CGCGTCAAGGTGAGGCGGTGGCGGAGCTTGGTCACTCATCTCAGATCGCGCTGCGAGA	1193	Db	1899	CGTGTCTGTGTGATCGATGCTGCTCTCCACACCAACCCACCATCTCTATCGAA	1958
Db	879	TAGAATTAATGATACACAGT-----AAATTGGTAATTAACAGATTAATTCCTGAAGA	932	Qy	2274	GTCAAAACAGCTCCACATTTCCATTTGGCTCGGTGGAAGAGCCAGACTTTTACCACTGTCAA	2333
Qy	1194	CGCTGGCTACTACAGTGGTGTAGCAGAAACAGCGCGGGAATCCCTGTGCGCTGCGCC	1253	Db	1959	GGCTAAACCTCATCTTCATCTTCTGCACTGGAGAGGCTGTGCAITTCACCGCTGACA	2018
Db	933	TGATGCTATTATCAGTGTGATGGTGAATAGCAAGGATCTATTTTATCTAGAGCCAG	992	Qy	2334	GATTTCAGCTACACTGTACGCTTCGGCCCTTGGGGGCTCAGGAATGCTTCCCTGGTCA	2393
Qy	1254	CCTGCGGTAGTGTGCGAGGCGCTGCCAGCGCCCGCTCGGCTCAGCGACCGCC	1313	Db	2019	AATCATTAATACTACCATCCGCTGTAATCTCTGTGTGCTGCAATGCTTCTTGTGTCT	2078
Db	993	ACTGACTGTAGTGTGAGAGACAGACCCAGTGTCTCTCTATTAATGTATGATGCTGAAC	1052	Qy	2394	CTACTATACAGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTA	2453
Qy	1314	GCTGAGCAGCTCTCTGTGTGGTGGCTGGAGCGGCTGAGTTGCAACGACGCAAAAT	1373	Db	2079	GTACCTTCAACAATCAGAACTCAGATGTTGGTTTCAAGGTCTAGAGGCTAGAAACCAATA	2138
Db	1053	CATGTCAAGCTCAGCCATCTTTTAGCTGGAGAGGCCACTTTTATAATTACAGACAAAGT	1112	Qy	2454	CGAGTTTGGGTACAGTCCACGAGTGGATATGGATGGGCTCTTGGCTCCGTCTGTAGA	2513
Qy	1374	CATTGGCTTCTCTCTACCAAAAGGCGAGGAGTGGACAAATGTGGAGTACCAAGTT	1433	Db	2139	CGAATTTGCGGTTCGATTAATATGGAATCAGCTTTCAGTCTCTGGAGCCCTGTAGTCTA	2198
Db	1113	CATTGGCTTATCTGTACACTACATCAAGCAGAGGTTTAAATTAATGAAGATATCAAGT	1172	Qy	2514	ACGCTCACCCCTGCTGACCGGCTTCAACACCTCTCTCTGCACTGCGCTGAGCCGCCCT	2573
Qy	1434	TGCAGTAAACATACACACAGAGCTGCAAGTTTCGGGACTGGAACCAACACGGAATTA	1493	Db	2199	CCATTCTACTTTCAGAGACCAAGCAGGCCCAACCACTGGAGTAAAGTGAATTAAT	2258
Db	1173	AGTCATCGAAATGACACACTCATTAATTAATTAATGATGACTTAGAGCTGCGCAAAITTA	1232	Qy	2574	GACACCATCCCGCTTGGGTGTTACATGTTGTGCTCCCAACGAGGCCCAATGGTGAATGT	2633
Qy	1494	TGAGTTCTACGTGGTGGCTACTCCAGCTGGGGCCAGCGCAACCTCCAGCCAGCCCT	1553	Db	2259	AGAGGATGACATCTGCTGCTGTTTGGAAACCCCTGTATGGCCCGAGAAACAGTTGTGAC	2318
Db	1233	TACTTCTACATTTAGCATATATGCCAATGGGAGCCAGCCAGATGTCTGACCATGTGAC	1292	Qy	2634	GGAGTATCTAATTTCTCTACAGCAACACACACCCAGCCGGAACCAAGTGGACATGCT	2693
Qy	1554	GGTGATACATGAGCAGATGCCAGCGCAGCAACCCAGCTTACCTTGTCCAGCCCA	1613	Db	2319	CCGCTATCTATCTTATATGCACTATGGAAGGCTTGATTTGAGGAGAGTGGCAGGCTT	2378
Db	1293	ACAGAACTACTAGAGGATGTTCCCTCTGAGACCTCTGAAATTTAGTTTGAACAATCGAAG	1352	Qy	2694	CACACAGAGGGAACATCTTCAAGTGCAGAGGTCCTAGGCTTAGAGAGTGCACACTCGGTA	2753
Qy	1614	CCCCTCGACATCAGGTTGGCATGCTGCCCTGCCCTCCAGCTGAGCAATGACAGGT	1673	Db	2379	ACACCGTGAAGGGCAATACCATGGCTTTGCTAGAAACTTGGTAGCAGGAATTTGTA	2438
Db	1353	TCCCACTGATATCTCATCTCTGCTGGCTGCCAATCCAGCCCAATATCGCGGGGCCAAGT	1412	Qy	2754	TTTCTTCAAGATGGAGCCCGCAGAGGTGGGCGCTTGGGCCCTTTTCCGCTTGCAGGA	2813
Qy	1674	GCTGAAGTACAAGATAGAGTACGCTTTGGGGAAGGAGATCAGGTTTTTCTCCACCGAGGT	1733				

Db 2439 CATTGTCAAGATATCTGCATCCAAATGAGTGGGAGAGGACCCCTTTTCABATTTCTGTGGA 2498
Qy 2814 TGTGATTAATCTGCGAAGAGACATTTCTCAGACTC 2846
Db 2499 GCTGGCAGTACTTCCAAAGGAAACCTCTGAATC 2531

RESULT 2

US-09-877-730-1
; Sequence 1, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3210
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-1

Query Match 5.1%; Score 316.6; DB 4; Length 3210;
Best Local Similarity 47.4%; Pred. No. 1.1e-71;
Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6;

Qy 300 AGACAGCATGATGAGGAAGCTTTAGGATCTGGAAGTCTACTGAGGCGAGCTATTCCTG 359
Db 279 ATACATCAGTGGTGGAGGCGAGGAGGAGCAGTCCGATGAAGGATTTATCAGTG 338
Qy 360 TCTGSCCCACAGCCGCTAGAGTGGTGGCCAGCCAGCTTGTGTTGCTCAGCTTGGCAC 419
Db 339 CTGGCAATGAACAAATAGAGGCAATTTCTAGTCAAAAGACTCATCTTGCCTTATCAAC 398
Qy 420 ACTCGAAGACTTCTCTGACACCCGAGTCCAGATTTGTGAGGAGAACGGGACAGCACG 479
Db 399 TATTCTGCATTTGAGTTCAGCCAAATTTCCACTGAGTCCAGAGTGGAGTTGCTCG 458
Qy 480 CTTTGAATGCCACCAAGGCGCTTCCAGCCCGCATCATTAATCTTGGGAAAGGACCAAGT 539
Db 459 ATTTGCATGCAAGATTTTCATCCACCCCTCTGCAGTCATAACATGGGAGTTCAATCGAC 518
Qy 540 GACCGTGCCTGAG---GAGCCCGGCTCATCTCTTCCAGTGGCTCTCCAGATCCT 596
Db 519 AACTCTACCTATGACTATGAGCAGGATACTGCCCCATCAACAGGAGTATTCAGATCTA 578
Qy 597 AGATGTCAGGACAGTGAAGTCTCTACCGTGGTGGCCACCAATTCAGCCCGCCA 656
Db 579 TGATGTCAGCAAGGGATCTGGAATATTCGTTGATTGCTGCCACTGTAGGCCACCG 638
Qy 657 ACGATTACGACGAGGCGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACCAG 716
Db 639 ACGTAAAGTATGGAGGCGCTCGCTAACTGTGATTCAG-----CTAAGGAGTCAAAATC 692
Qy 717 GGGGAGGATGTTGATTTGGCGAGCCCGAGAGAACACCAAGTATGTTCTGGAAGAGAA 776
Db 693 CTTCCACACACCAATATAGCAGGTCCACAGAAACATAACAACATCTCTTTCATCAGAC 752
Qy 777 TGATGATGAGTGGGCTCTGCTGACCCCGCTTTTGTCTCGGTGGGTCCGACA 836
Db 753 TGTAGTTTGGATGCTGGCCACAGGAAATCCCAAAACCAATCATTTCTTGGAGCCGCT 812

Qy 837 GGATGGAAGCGCTAT---CTCCACGGATGTCTCGTTCTGGCGCGGACCAATCTACTCAT 893
Db 813 TGATCACAAAATCCATTTGATGTCTTTAATACTCGGGTACTTTGAAAATGTTAATCTCATGAT 872
Qy 894 CGCCAGCGCGCAGCCTCGGCACCTCTGGAGTCTATGTCTGCGGAGCCCAACAGCCCTCCAC 953
Db 873 ATCTGATGTCAGGCTACAAATGCTGAGTATATGTTTGTTCGGGCCACTACCCCTGGCAC 932
Qy 954 GCGTGACTTCCCACTCGCGCTGCTCGAGTCTCGAGTCTCGAGTCTTGTCTCCCGCAGCATCTCGCA 1013
Db 933 ACGCAACTTTACAGTTGCTATGCGCACTTTAACTGTATTAAGTCTCTCTCTCATTTGTTGA 992
Qy 1014 GGCACCGAGGCGCTCTCGGAGCGGGGCCAGGACCGCGGCTTCTGCTGCGCGGGCTC 1073
Db 993 ATGGCCAGAAAGTTTAAACAAGGCGCTCGAGCTGGCACTGCTCGATTTGTGTGTGAGGAGCA 1052
Qy 1074 CGGGAGCCACGCGCCCGCTGCACTGCTGCGACGAGGATCCGTTTGGAGCCCAATGG 1133
Db 1053 AGGAATCCCTCTCTCCCAAGATGTCTATGTTGMAAATGGAAGAGATACATTCGATGG 1112
Qy 1134 GCGCGTCAAGGTGACGGCGGTGGCGCACTTGGTCTATCATCTCAGATCGGCTTGCAGGA 1193
Db 1113 TAGAATTTAAAATGTACAAACAGT-----AAATGGTAAATTAACCCAGATTTATCTCGAAGA 1166
Qy 1194 CGCTGGCTACTACCAAGTGGCTAGCAGAAACAGCGCGGGAACCTGCTGTGCGCGTGGGCC 1253
Db 1167 TGATGCTATTTATCAGTGCATGCTGAGAAATAGCAAGGATCTATTTATCTAGAGCCAG 1226
Qy 1254 CTTGGCGGTAGTGTGCGCGAGGGGCTGCCAGCGCCCGACTCGGGTTCACAGCCAGCCGCC 1313
Db 1227 ACTGACTGTGATGTCTCAGAAACAGACCCAGTGTCTCTCTATTAATGTATCATGTCTGAAC 1286
Qy 1314 GCTGAGCAGCTCTCTGCTGTGTGGCTTGGAGCGGCTGAGTTGTCACAGCGAGCAAT 1373
Db 1287 CATGCTCAAGCTCAGCCATCTCTTTAGCTTGGAGAGGCCACTTTTATTAATTCAGACAAAGT 1346
Qy 1374 CATTTGGCTTCTCTTCTCACTACAAAGGACAGGGGAGTGACAAATGTGAGATACCAAGT 1433
Db 1347 CATTTGCCCTATTTCTGTACACTACATGAAGCAGAGGTTTAAATTAATGAAGATATCAAGT 1406
Qy 1434 TGCAGTAAACAAATGACACACACAGCTGCGAGTTGGGAGCTGGAACTGGAACCCACACGGATTA 1493
Db 1407 AGTCATCGGAAATGACACAACTCATTAATTAATTAATGATGACTTGAAGCTTGCAGCAATTA 1466
Qy 1494 TGAGTTCTAGCTGTGGCTTACTCCAGCTGGGGGCGAGCGCACTCCAGCCCGAGCCCT 1553
Db 1467 TACTTTCTCATTTGTAGCATATATGCCAATGGGAGCCAGCCAGATGTCTGACCATGTGAC 1526
Qy 1554 GGTGCATACACTGGAACGATGTCCCGAGCGGAGCAACCCAGCTTACCTTGTCCAGCCCAA 1613
Db 1527 ACAGAAATCTCTAGAGGATGTTCCCTGAGACCTCTCTGAAATTAAGTTTGAACAAGTCGAAG 1586
Qy 1614 CCCTCGGACATCAGGGTGGCATGGCTGCGCCCTGCTCCAGCTTACGAGCAATGAGCAGGT 1673
Db 1587 TCCACATGATTTCTCTCATCTCTGCTGCTGCAATCCCAAGCCCAATATCGGGGGGCCAAGT 1646
Qy 1674 GCTGAAGTCAAGATAGAGTACGTTTGGGGAAGGAGATCAGGTTTCTCCACCGAGGT 1733
Db 1647 GGTGCTGTATGCTGTCTTTCGCTTAAGTACTGAGAAATTCATCCAGTTCTGAGAGCT 1706
Qy 1734 GCCTGGAATGAGACAACTTACGTTTAACTCACTTTCAGCCCAACAAAGTGTACCGAGT 1793
Db 1707 CCGGGGACCAACATGAGTACCTTTTGGAAAGCCCTGAAACCTGACAGTGTCTACCTGGT 1766
Qy 1794 CCGGATTTACGCTGGCATCTGGCTGATGAGTCCCTTCTCAGTGGATGCAAGCAGACAG 1853
Db 1767 TCGGATTTACTGCTGCCACAGAGTGGGGCTGGGAGTGCATCAGTATGGACTTCACATAG 1826
Qy 1854 GACACCTGGTGTGCACAAACAGAGCCATGTTTCCCTTTGGCCCTTGCAGAAATGAGGTGAG 1913
Db 1827 GAGCCCAAGCTTACAAAGCTGAAAGCCCTTAAGTCTCCAGAGTTGCAITTTGGAGCTCT 1886
Qy 1914 GGCAAGATGGAGTCCCTGCTGTGTGTCTATGCGAGCCGCCCTCCACCCACCCAGATCTC 1973

1887 GAACGTACCAACCAATTTCTGAGGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCA 1946
1974 TGATACAAACTCTACTTGGGAGAGTGGGAAACAGAGGAGGAGGAGAGTGGTACGCCCC 2033
1947 GGGCTACAAAGCTGTACTACAAAGAAAGGAGGAGGAGAA-----1988
2034 CCAGGGGGTCTGTGGAGATCAAGCTTTGGGAGCTCGGGCCCGTGGCTGGAAGAAGAGT 2093
1989 -----TGGGCCCACTTTCTTGGATACCAAGGA 2015
2094 GAAGCAGATGAACTGACCAAGTGTAGTCCCTGGCAGGCCGTGACAGAGTGAAGTCTGTAGC 2153
2016 CTTACTCTACTCTCTAGTGGCTTAGACCCCAAGAAATATATCTGTGAGACTCTCTGGC 2075
2154 TTTCACAAACACAGAGAGGCTACGCTGTGTGTGAAAGGCAAGCAGGAGGCGCC 2213
2076 TTACAAACAATAGACATGGCTATCAGGAGATCAGACTGTTCAGCAC---TCCAGGATG 2132
2214 CAGCCAGAGCTTCCTATCAGAGGGGGCCACCGCTCCCTTCCCATGTTCACCGCAGA 2273
2133 CGTGTCTGTTCGTGATGCGATGCTCCTCTCCACACACCCACCACTCTCTATGGAA 2192
2274 GTCAACAGCTCCACCTTCATTTGGCTTGGTGGAAAGAGCCAGACTTTTACCACTGTCAA 2333
2193 GGCTAACACCTCATCTTCCATCTTCTGCACTGGAGAGGCTGCACTTCAACGCTGCACA 2252
2334 GATTGTCAACTACACTGTACGCTTGGCCCTGGGGCTCAGAAATGCTTCCCTGTCTAC 2393
2253 AATCATTAACATACCACTCCGCTGTAACTCCTGTGTGGCTTGCAGAAATGCTTCTTGTCT 2312
2394 CTACTATACCAAGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCACTTTTACCAAGTA 2453
2313 GTACCTTCAACATCAGAACTCAGATGTTGGTTCAGGTCTAGAACCAACCACTAATA 2372
2454 CGAGTTTGGGGTACAGTCCCAAGAGTGGATATGAGTGGGCTTGTGGCTTCGTGTAGA 2513
2373 CGAATTTGGCGTTTGGATTAATGATGATCAGCTTTCAGTCTTGGAGCCCTGTAGTCTA 2432
2514 ACCTCCAGCTGCTGACCGGCTTCAACCTCTCTGACCTGGGCTGAGCCCT 2573
2433 CCATTTCTACTTCTCAGAGAGCACCAGCGGCCCACTGTTGGAGTAAAGTGAATTAAT 2492
2574 GACACCATCACCGTTCTGGTTACATGTTGTCCTCCCGAGGCCCAATGCTGAGATTGT 2633
2493 AGAGGATGACCTGCTGCTGTTTCTTGGAAACCCCTGTAGTGGCCAGAAACAGTTGTAC 2552
2634 GAGATATTAATTTCTACAGCAACCAACCAACCCAGCCCGAAACACAGTGGACACTGTCT 2693
2553 CCGCTACTATCTTATATGCACTAGGAAGGCTGGATTTGCAGGAGAGTGGCAGGCTCT 2612
2694 CACCAAGAGGAAACATCTTCTAGTGCAGAGGTCATGGCTAGAGAGTGAACACTCGGTA 2753
2613 ACACCGTGAAGGGGCAATAACCACTGCTTGTAGAAACCTTGGTAGCAGGAAATGTGA 2672
2754 TTCTTCAAGATGGGAGCCGCAAGAGTGGGCGCTGGGCTTGTTCGCGCTTTCAGGGA 2813
2673 CATTTGACAGATCTGATCTCAATGAGTGGGAGAGGACCCCTTTTCAAAATTTCTGTGA 2732
2814 TGTGATTACTCTGCAAGAGACATTTCTCAGACTC 2846
2733 GCTGGCAGTACTTCCAAAGGAACCTCTGAATC 2765

RESULT 3
US-09-877-730-17
; Sequence 17, Application US/0987730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632e1 Human Phosphatases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3219
; TYPE: DNA
; ORGANISM: homo sapiens
; US-09-877-730-17

Query Match 5.1%; Score 316.6; DB 4; Length 3219;
Best Local Similarity 47.4%; Pred. No. 1.1e-71;
Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6;
QY 300 ACACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCG 359
DB 45 ATACATCAGTGAAGTGAAGGCGGAGGAGAGCAGTCCGATGAAGGATTTTATCAGTG 104
QY 360 TCTGGCCACACCCGCTAGGAGTGGTGGCCAGCCAGGTTGTGTGGTCAAGCTTGCAC 419
DB 105 CTTGGCAATGAACAAATATATGGAGCCATTCTTAGTCAAAAAGCTCATCTTGCCTTATCAAC 164
QY 420 ACTCGAAGACTTCTCTCTGACCCGAGTCCAGATTTGTGGAGGAGAACGGGACAGCAG 479
DB 165 TATTTCTGCATTTGAAGTCCAGCCAAATTTCCACTGAGGTCCACGAAGTGGAGTTGCTCG 224
QY 480 CTTTGAATGCCACACAGGGGCTTCCAGCCCATCATTTACTTGGGAAAGGACCGAGT 539
DB 225 ATTTGATCAAGATTTTCATCCACCTCTCGAGTCATAACATGGAGTTCATTCGGAC 284
QY 540 GACCGTCCCTGAG---GAGCCCGGCTCATCACTCTTCCAAAGTGGCTCTCCAGATCCT 596
DB 285 AACTCTACTATGACTATGGAAGATTAAGTCTGCTTACCAACAGGATTTATGAGATCTA 344
QY 597 AGATGTCCAGGACAGTGAATGAGGCTCTTACCGCTGGCGCCCAACATTCAGCCGCCA 656
DB 345 TGATGTCCAGCAAAAGGATTTCTGGAATTTATCGTTGTTATTTGCTGCACTGTAGCCACCG 404
QY 657 AGATTTCCAGCAGAGGCTCGCTCACTGTGGCCCTCAGAGGCTCTTGGAGGCTACCA 716
DB 405 ACGTAAAGATGGAGGCTCTGCTAACTGTGATTCAG-----CTAAGGAGTCAAAATC 458
QY 717 GGGGCAAGATGTGGTCATTTGGCAGCCCCAGAGAACACACAGGTAGTGTCTGGACAGAA 776
DB 459 CTTCCACACACCAACATTTATAGAGGTCACAGAACATAACACATCTCTTCATCAGAC 518
QY 777 TGTAGTGAATGAGTGGCTGTGCTGACCCCAACCCCTTTTGTGTCTGTGGTCCGACA 836
DB 519 TGTAGTTTGGAAATGATGCGCCACAGAAATCCCAACCAATCATTTCTTGGAGCCGCT 578
QY 837 GGATGGAAGCCATAT---CTCCACGATGTCTATCGTTCTGGGCGGACCAATCTACTCAT 893
DB 579 TGATCAAAATCCATTTGATGTCTTTAATATCTCGGGTACTTGGAAATGGTAATCTCATGAT 638
QY 894 CGCAGCGCGCAGGCTCGGCACCTCTGAGTCTATGTCTCGCAGGCAACAAAGCCCTCAC 953
DB 639 ATCTGATGTCAGGCTACAAACATGTGGAGTATATGTTGTGGGCCCATCTACCCCTGGCAC 698
QY 954 GCGTGAATTCGCCACTGCGGCTGTGAGCTCCGAGTGTGTGTGGCCCAAGCCATCTCGCA 1013
DB 699 ACGCAACTTTACAGTTGCTATGCAACTTTTAACTGTATTAGCTCTCTCTTCAATTTGTTGA 758
QY 1014 GGCACCCGAGGCGCTCTCGCGACCGGCGCCAGCAGCGGCGCTCTGTGTGCCGGCGTC 1073
DB 759 ATGGCCAGAAAGTTTAAACAAAGGCTCGAGCTGCACTGTCTCGATTTGTGTGTCAGGCAGA 818

Qy	1074	CGGGAGCCACGGCCCGCGCTGCACTGGCTGCAAGACGGGATCCGTTGCGAACCCAAATGG	1133
Db	819	AGGAATCCCTCTCCCAAGATGTCAATGGTTGAAAAATGGAAGGAAGATACATTCGAATGG	878
Qy	1134	GC CGCTCAAGGTGCAGGGCGGTGGCGGACGCTTGGTCATCACTCAGATCGCCTGCAGGA	1193
Db	879	TAGAAITAAATGTACAACAGT-----AAATTGGTAAATTAACCCAGATTAATTCCTGAAGA	932
Qy	1194	CGCTGGCTACTACCAAGTCGGTAGCAGAAAAACAGCGCGGAACCTGCCCTGTGCTGCGCC	1253
Db	933	TGATGCTATTATCAAGTGCATGGCTGAGATAGCCAAGGATCTATTTTATCTAGAGCCAG	992
Qy	1254	CTTGGCGGTAGTGTGGCGAGGAGGTGCCCCAGCGCCCCGAGCTCGGGTCAAGCACGCC	1313
Db	993	ACTGACTGTAGTGAATGTGAGAAGACAGACCAGTGTCTCCCTATAATGTACATGCTGAAC	1052
Qy	1314	GCTGAGCAGCTCCTCTGCTGTGGTGGGCTGGGAGCGGCTGTGATGTCACAGCGAGCAAT	1373
Db	1053	CATGTCGAAGCTCAGCCAAATCTTTTAGGCTGGGAGGCCCACCTTTATTAATTCAGACAAAGT	1112
Qy	1374	CATTGGCTTCTCTCTTCACTACCAAAAGGCGAAGGGAGTGACAAATGTGAGGTACCAAGTT	1433
Db	1113	CATTGCCATTTCTGTACACTACATGAAGCAGAGGTTTAAATATGAAGGATATCAAGT	1172
Qy	1434	TGCAGTAAACAAATGACACACAGAGCTGCAGGTTGCGGACCTGGAACCCCAACACGGATTA	1493
Db	1173	AGTCATCGGAAATGACACAACTCAATTATATTAATGATGACTTAGAGCTGCGCAGCAATTA	1232
Qy	1494	TGAGTTTCAAGTGTGGCTTACTCCAGCTGGGGGCGAGCGGAACCTTCAGCCCGCCCT	1553
Db	1233	TACTTTTCTACATGTAGCATATATGCCAAATGGGAGCCAGCCAGATGTCACCATGTGAC	1292
Qy	1554	GGTGACATCACTGGACGATGTCCCGACGCGACACCCACAGCTTACCTTGTCTCAGCCCCAA	1613
Db	1293	ACAGAACTCTAGAGGAATGTTCCCTGTGACCTCTGAAATTAAGTTTGAAGTGGAG	1352
Qy	1614	CCCTCGGACATCAGGGTGGCATGGCTGCCCCCTGCCCTCCAGCCTGAGCAATGACACAGGT	1673
Db	1353	TCCCACTGATATTCTCATCTCTGGCTGCCAAATCCAGCCAAATATCGCGGGGCCAAGT	1412
Qy	1674	GCTGAAGTACAGATAGAGTACGGTTTGGGGAGGAAGATCAGGTTTTCTCCACCGAGGT	1733
Db	1413	GGTGCTGTATCGCTGTCTTTTCGCGCTAAAGTACTAGAAATTCAAATCCAAAGTCTGGAGCT	1472
Qy	1734	GCCTGGAATGAGACACAACTTACGTTAAACTCACTTCAGCCAAAACAAAGTGATCCAGGT	1793
Db	1473	CCCGGGACCAAGCATGAGTACCTTTTGGAAAGCGCTGAAACCTGACAGTGTCTACCTGGT	1532
Qy	1794	CCGATTTCACTGGCACTGGCGTGGCTATGAGAGTCCCTTCTCAGTGGATGACGACAG	1853
Db	1533	TCGGATTACTGCTGCCACAGACAGAGTGGGGCTGGGAGAGTCATCAGTATGGACTTCCACATAG	1592
Qy	1854	GACACCTGGTGTGCACAAACACAGACCATGTTCCCTTTGGCCCTGCGAATTTGAAGGTGAG	1913
Db	1593	GAGCCCAAAGCTACAAGCGGTGAAGCCCCCTAAAGTCTCCAGAGTTGCAATTTGGAGCCTCT	1652
Qy	1914	GGCAAAAGATGGAGTCCCTGTGTGGTGTCAATGGCAGCCGCCCTCCACCCACCCAGATCTC	1973
Db	1653	GAACTGTACCAACCATTTCTGTGAGTGGCAGCAAGATGTGAGAGCACAGCTGCTATTCA	1712
Qy	1974	TGGATCAAACTCTACTGGGGAGAGGTGGAAACAGAGGAGGAGGAGATGGTGACCGCCCC	2033
Db	1713	GGGCTACAAGCTGTACTACAAGGAAGAAGGGCAGCAGAGAA-----	1754
Qy	2034	CCCGGGGGTGTGGAGATCAAGCTTGGGAGCTCGGGCGCGGTGGAAGGAAGAAAGT	2093
Db	1755	-----TGGGCCCAATTTCTTGTGATACCAAGGA	1781
Qy	2094	GAAGCAGTATGAAGTCAACCAATTAGTCCCTGGCAGCGCTGACGAGTGAAGCTCGTAGC	2153
Db	1782	CCTACTCTATCTCTAGTGGCTTAGACCCCAAGAAAGAAATATCATGTGAGATCTCTGGC	1841
Qy	2154	TTTCAACAAACACGAGGACGGCTACGCTGTGTGTGGAAAGGGCAAGACGAGGAAGCGCC	2213

RESULT 4

RESULT 4
IIS-09-977-730-31

US-09-877-31
: Sequence 31 Application IIS/09877730; sequence 31, Appl
: Patent No. 6465632

; PATENT NO. 6463632
;
; GENERAL INFORMATION:

APPLICANT: Walke, D. Wa

APPLICANT: WAJKE, D. WAJKE
APPLICANT: Scoville, John

APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

APPLICANT: Abuin, Alejandro

APPLICANT: Zambrowicz, Brian

APPLICANT: Sands, Arthur T.

TITLE OF INVENTION: No. 646566

FILE REFERENCE: LEX-0189-USA

; CURRENT APPLICATION NUMBER: US/09/877,730

; CURRENT FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 60/210,607

; PRIOR FILING DATE: 2000-06-09

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Fas

; SEQ ID NO 31

; LENGTH: 3

; TYPE: DNA

; ORGANISM: ho
118-00-977-730-31

[illegible]

RESULT 5

```

US-09-877-730-7
; Sequence 7, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3453
; TYPE: DNA
; ORGANISM: homo sapiens
; US-09-877-730-7

```

	Query Match	5.1%;	Score 315.4;	DB 4;	Length 3453;
	Best Local Similarity	47.4%;	Pred. No. 2.3e-71;		
	Matches 1210;	Conservative	0;	Mismatches 1271;	Indels 72; Gaps 6;
Qy	300	AGACAGCGATGATGAGGAAGCTCTTAGGNTCTGGAAGTCACTGAGGGCAGCTATTCTGT	359		
Db	279	ATACATCAGTGAAGTGAAGCGACGGCAGGAGAGCAGTCCGATGAAGGATTTTATTCAGTG	338		
Qy	360	TCTGGCCACAGCCCGCTAGAGTGTGTGGCCAGCCAGGTCTGTGCTCAAGCTTTGGCAC	419		
Db	339	CTTGGCAATGACAAATATGGAGCCATCTTAGTCAAAAGCTCATCTTGCCTTATCAAC	398		
Qy	420	ACTCGAAGACTTCTCTCTGCAACCCGAGTCCCAGATTGTGGAGGAGAACCGGACAGCACG	479		
Db	399	TATTTCTGCATTTGAAGTCCAGCCAAATTTCCACTGAGGTCCACGAAGGTGGAGTTGTCTG	458		

QY	480	CTTTGAATGCCA	CHACCAAGGGCTT	CCAGCCCCCATCATTA	CTTACCTTTGGGAAAGGACCAAGT	539
DB	459	ATTTCGATGCAAGATTTT	CATCCACCCCTCTCGAGTCA	TAACTAATGGGAGTTCAATCGGAC	518	
QY	540	GACCGTGCCTGAG--	--GAGCCCCGGGCTCATCA	CTCTCCCAAGTGGCTCTCCAGATCCT	596	
DB	519	AACCTCTACCTATGACTAT	TGGACAGGATAA	CTGCCCTTACCACAGGAGTAT	TTCAGATCTA	578
QY	597	AGATGTCCAGGACAGT	GATCAGGCTCCTAC	CGCTGGTGGCCACCACTAAT	CAGCCCGCCCA	656
DB	579	TGATGT	CAGCCAAAGGATTTCT	GGAAATTAATCGTTGTA	TTCGCTGCGCACTGTAGCCCAACG	638
QY	657	ACGATTCAGCCAGGAGCGCT	CGCTCACTGTG	GGCCCTCAGAGGGTCTTTTGAGGCTAC	CCAG	716
DB	639	ACGTAAAGAT	TGGAGGCTCGCTAA	CTGTGAATCCAG-----	CTAAGGAGTCAAAATC	692
QY	717	GGGCGAGGATGTGT	CAATCTGTGCAGCC	CCAGAGAACACCA	CGGTAGTGTCTGGACAGAA	776
DB	693	CTTCCACACACCA	CAATATATAGCAGGT	CCACAGAACATAACAA	CACTCTTTCATCAGAC	752
QY	777	TGTAGTGATGGAGT	CGCTGGCTCTG	CTGACCCCAACCCCTTTTG	TGTCTGGGTCCGACA	836
DB	753	TGTAGTTT	TGGATGATGCGCA	CAGGAATCCAAACCAATCAT	TTCTTGGAGCGCCT	812
QY	837	GGATGGAAAGCTAT--	---CTCCACGGATGCA	TGTTCTGGGCGGACCAATCT	ACTCAT	893
DB	813	TGATCACAAT	CCAATTGATGCTTT	TAATACTCGGCTACTTGG	AAATGGTAATCTCATGAT	872
QY	894	CGCAGCGCGCAGCCT	CGGCACCTCTG	AGATCTATGCTGCGAGCCAA	CAGCCCTCAC	953
DB	873	ATCTGATGT	CAGGCTACA	CAATGCTGAGTATATGTTGT	CGGCGCCACTACCCCTGGCAC	932
QY	954	GGGTGACTT	CGCCACTGCGGCTG	CTGAGCTCCGAGTCTTGCT	TGCCCCAGCACTCGCA	1013
DB	933	ACGCAACTTT	CAGTTGCTATGGCA	CTTTAACTGTAT	TAGTCTCTCTTCATTTGTTGA	992
QY	1014	GGCACCGAGGCGCT	CTCGCGGACCGCGGCGCAG	CACCGCGCTTCGTGTGCGGGCGCTC	1073	
DB	993	ATGGCCAGAAAGTTT	TAAACAGGCTCGAGCT	TGGCACTGCTCGATTTGTGTGTCAGGCAGA	1052	
QY	1074	CGGGAGCCACGGCCCGCT	CGACTGGCTGACAGCAGGAGAT	CCGTTTGGGACCCAAATGG	1133	
DB	1053	AGGAATCCCTCT	CTCCCAAGATGTCA	TGTTTGAATAATGGAAGGAGATACATA	TCGAAATGG	1112
QY	1134	GCGCGTCAAGTGCAGGCGG	GTGGCGCAGCTTGCTG	TCATCACTCAGATCGGCTG	CGGACGA	1193
DB	1113	TAGAAATTA	AAATGTACAACAGT-----	AAATTTGTAATTAAC	CAGATTTATTCCTGAAGA	1166
QY	1194	CGCTGGCTCATCTAC	GATGCGGTAGCAGAA	AAACAGCGCGGAACTGCCGTGCGCGTCGCC	1253	
DB	1167	TGATGCTATTAT	TCAGTGCATGGCTGAGAT	TAGCAAGGATCTATTTTATCTAGAGCCAG	1226	
QY	1254	CTTGGCGGTATGTGTG	CGGAGGGCTG	CCGAGCCCGGACTCGGGTTC	CAGCACGCC	1313
DB	1227	ACTGACTGTATGTGT	CAGAAGACACAC	CCAGTGTCTCCCTATAA	TGTACATGCTGAAC	1286
QY	1314	GCTGAGCAGCTCCT	CTGTGTGTGGCTCGGAGCGGCT	TGATGTGACAGCGAGCAAT	1373	
DB	1287	CATGTCAAGCT	CAGCCAAATCTTTT	TAGCCTGGGAGGCGCACTTTATTAATTCAGACAAAGT	1346	
QY	1374	CATTGGCTTCTCT	CTTCACTACCAAAAGGCA	AGGGAGTGGCAATGTGAGGTACCAAGT	1433	
DB	1347	CATTGCTTAT	TCTGTACACTACATGA	AAAGCAGAGGTTTAAATTAATGAAGTATCAAGT	1406	
QY	1434	TGCAGTAAACAA	TGACACCA	CAGAGCTG	CAGGTTTCGGGACCTCGGAAACCAACGAGTTA	1493
DB	1407	AGTCATCGGAAAT	TGACACAACTCAT	TATATTAATTAATGA	CTTAGAGCCTGCGACCAATTA	1466
QY	1494	TGAGTTCT	CATCGTGGCGCTACT	CCAGCTGGGGGCGCAGCGGAACCTC	CAGGCCAGCCCT	1553
DB	1467	TACTTTCT	CATATGTAGCATATAT	GCCAATGGAAGCAGCAGATGCT	GTGACCATGTGAC	1526

```
QY 1554 GGTGCATACACTGGACGATGTCCCGAGCGCAGCACCCTTACCTTGTCTCAGCCCA 1613
D 1527 ACAGATACTCTAGAGGATGTTCCCTGAGACCTCTCTGAAATTAGTTTGACAAAGTGA 1586
QY 1614 CCGCTCGGACATCAGGGTGGATGCTGCTGCCCTGCTCCCTCAGCCTGAGCAATGGACAGT 1673
D 1587 TCCCACTGATATTTCTCATCTCTGCTGGCTGCAATCCAGCAAAATATCGCGGGGCAAGT 1646
QY 1674 GCTGAAGTACAAGATAGATGACGTTTGGGGAAGGAAGATCAGGTTTTCTCCACCGAGT 1733
D 1647 GGTGCTGTATCGTGTCTTTCCGCTTAAGTACTGAGAAATCAATCCAAAGTTCTGAGCT 1706
QY 1734 GCTGGAAATGAGACAACTTACGTTAACTCACTTCAGCCCAAAAGTGTACCGAGT 1793
D 1707 CCGGGGCAACGCGATGATACCTTTTGAAGGCGCTGAAACCTGACAGTGTCTACTCTGT 1766
QY 1794 CCGGATTCAGCTGGCACTGGCGCTGCTATGAGTCCCTTCTCAGTGGATGACGACAG 1853
D 1767 TCGGATTACTGCTGCCACAGAGTGGGCTGGGAGAGTCAATCAGTATGAGCTTCAATAG 1826
QY 1854 GACACCTGCTGTCACAAACAGAGCCATGTTCCCTTTGCCCCCTGCAGAAATTCAGGTGAG 1913
D 1827 GACGCCAAAGCTACAAGGTGAAGCCCTTAAGTCTCAGAGTTGCAATTTGGAGCTCT 1886
QY 1914 GGCMAAGATGGAATCCCTCTGCTGTGTATGGCAGCGCCCTCTCACTCCACCCAGATCTC 1973
D 1887 GAACCTGACCACTTTCTGTGAGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCA 1946
QY 1974 TGGATCAAACTTACTTGGGAGAGTGGAAACAGAGGAGGAGGAGATGTTGACCGGCC 2033
D 1947 GGGCTACAAGCTGTACTACAAGGAAGAGGCGAGCAGAGAA 1988
QY 2034 CCAGGGGTCTGGAGATCAAGCTTGGGACGTGGGCGCGTGGGCTGAAGAAAGT 2093
D 1989 -----TGGGCGCAATTTCTTGGATACCAAGA 2015
QY 2094 GAAGCAGTATGAATCAGCCAGTTAGTCCCTGGCAGCGCTACGAGTGAAGTCTGTAGC 2153
D 2016 CTTACTCTATCTCTCAGTGGCTTAGACCCAGAGAAATATCATGTGAGACTCTCTGC 2075
QY 2154 TTTCAACAAACAGGAGCGGTACCGCTGTGTGTGTGAAGGGGCAAGCGAGAGCGGCC 2213
D 2076 TTACAACAAATAGACAGATGGCTATCAGGCAGATCAGACTGTGAGCAC 2132
QY 2214 CAGCCAGACCTGCCCTATCCAGAGGGGCGCACCGCTGCTCTGCGCAATGTCACGCGA 2273
D 2133 CGTGTCTGTGTGATGCGATGCTCCTCTCCACCAACCCACCCACCTCTCTATGCGAA 2192
QY 2274 GTCAACAGCTCCACTTTCATTTGGCTTGGTGGAGAGAGCCAGACTTTACCACTGTCAA 2333
D 2193 GGGTAACACTCATCTTCATCTTCTGCACTGGAGAGGCTGCAATTCACCGCTGCACA 2252
QY 2334 GATTTGCAACTACACTGTACGCTTGGCCCTTGGGGCTCAGGAATGCTTCCCTGCTAC 2393
D 2253 AATCAATTAACATACACCATCCGCTGTAATCTCTGTGGCTGCAAGATGCTTCTTGTGTTCT 2312
QY 2394 CTACTATACGCTCTGGHAGACATTTCTCATTTGGCGGCTGAACACCAATTTACCAAGTA 2453
D 2313 GTACCTTCAACACATCAGAACTCACATGTTGGTTCAGAGTCTAGAACCAACCAAAATA 2372
QY 2454 CGAGTTTGGGTACAGTCCCAAGAGTGGATATGGATGGCGCTTTGGCTCGGTGAGA 2513
D 2373 CGAATTTGGCTGCAATTAATATGTGATCAGCTTTCAGTCTTGGAGCCCTGTAGTCTA 2432
QY 2514 AGCTCCACCTGCTGACCGGCTTTCAACACTCTCTTGAACCTGCGCTGAGCCCTCT 2573
D 2433 CCATTTCTACTTCCAGAGACACAGCAGGCCCCACAGTTGGAGTAAAGTGAATTAAT 2492
QY 2574 GACACCATCCACCGTTCGTTACACTGTTGTGTCGCCCAAGAGCCCAATGGTGAATGTT 2633
D 2493 AGAGGATGACACTGCTGCTGTTCTTTGGAAACCCCTGATGGCCCAAGAAACAGTTGTGAC 2552
QY 2634 GGAGTATCAATTTCTCTCAGGCAACAAACCAACCCAGCCCGAACCAGTGGACACTGCT 2693
```

```
D 2553 CCGCTATACTATCTTATATGCAATCTAGGAAGCCCTGATTCAGAGAGAGTGGCAGGCTT 2612
QY 2694 CACCACAGAGGAAACATCTTCACTGAGAGGTCATAGGCTTCAAGCTAGAGTGAATCTGCTA 2753
D 2613 ACACCGTGAAGGGCAATAACATGCGCTTCTAGAAAACCTTGGTAGCAGGAATGTGTA 2672
QY 2754 TTTCTTCAAGATGGAGCCCGCACAGAGTGGGCGCTTGGGCGCTTTTCCCGCTTGCAGGA 2813
D 2673 CATTTCAAGATATCTGCAATCAATGAGTGGAGAGGAGCCCTTTTCAAAATTTCTGTGA 2732
QY 2814 TGTGATTACTCTGCAAGAGACATTTCTCAGACTC 2846
D 2733 CTTGCGAGTACTTCCAAAGGAAACCTCTGAATC 2765
```

RESULT 6

```
US-09-877-730-15
; Sequence 15, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-15
```

```
Query Match 3.5%; Score 217; DB 4; Length 2481;
Best Local Similarity 50.3%; Pred. No. 5.8e-46;
Matches 647; Conservative 1; Mismatches 621; Indels 18; Gaps 4;
```

```
QY 300 AGACAGCATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGCGAGCTATTCCTG 359
D 45 ATACATCAGTGAAGTGGAGGCGAGGCGAGGAGAGCAGTCCGATGAAGGATTTTATCAGTG 104
QY 360 TCTGCCCCACAGCCCGCTAGGAGTGGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGGCCAC 419
D 105 CTGGCAATGAACAAATATGGAGCCATTTCTTAGTCAAAAAGCTCATCTTGCCTTATCAAC 164
QY 420 ACTCGAAGACTTCTCTGCAACCCCGAGTCCCGAGTCTGGAGGAGAGACGGGACAGCAG 479
D 165 TATTTCTGCAATTTGAATCCAGCCAAATTTCCACTGAGTCCCAAGAGTGGAGTGGCTG 224
QY 480 CTTTGAATGCCACCAAGGGGCTTTCAGGCCCCCATCAATTAATTGGGAAAAAGGACAGGT 539
D 225 ATTTGATGCAAGATTTTCATCCACCTCTCTGCACTATAACATGGGAGTTCAATCGGAC 284
QY 540 GACCGTGCCTGAG---GAGCCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCT 596
D 285 AACTCTACTATGACTATGGAAGGATTAAGTAACTGCGCTACCAAGAGGATTTGCAAGTCTA 344
QY 597 AGATGTCCAGGACAGTGAATGAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCA 656
D 345 TGAATGACCCAAAGGATTTCTGGAATATGTTGTTATGTTGTCACATGTAGCCCAAG 404
QY 657 AGGATTCAGCCAGGAGGCTCTGCTCACTGTGCGCCCTCAGAGGCTCTTTGGAGGCTTACAG 716
D 405 AGTAAAGTATGAGGCGCTCGCTAACTGTG-----ATTCCAGCTAAGGAGTCAAAATC 458
```

```
717 GGGGAGGATGCTGCTCATTTGGGAGCCCGGAGAAACACCGAGTGTCTGACAGAA 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
459 CTTCCACACACCAATATAGCAGGTCACAGAACATACACATCTCTTCATCAGAC 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
777 TGTAGTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
519 TGTAGTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
837 GGATGAAAGCTAT--CTCCAGGATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 893
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
579 TGATCACAATCCATTTGATGCTTTAATCTCGGCTATCTGGAATGCTATCTCATGAT 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
894 CGCCAGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
639 ATCTGATGCTAGGCTTACACATCTGAGATATATGTTGTCGGGCACTACCCCTGGCAC 698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
954 GGTGACTTGGCACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1013
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
699 ACGCAACTTTACAGTTGCTATGGCACTTTAACTGCTATGCTGCTGCTGCTGCTGCTG 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1014 GGCACCCGAGGCTCTCGCGAGCGCGGCGGAGCAACCGGCTGCTGCTGCTGCTGCTGCTG 1073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
759 ATGGCCAGAAAGTTAAAGGCTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1074 CGGGAGCCAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
819 AGGAATCCCTCTCCAGATGCTCATGTTGAAATGGAAGGATACATTCGATGG 878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1134 CGCGTCAAGGTGCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
879 TAGAATTAATTAATGTAACAGT-----AAATGCTAATTAACAGATTTATTCCTGAAAGA 932
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1194 CGCTGCTACTACAGTGGTAGAGCAAAACAGCGGGAAGTCTGCTGCTGCTGCTGCTGCTG 1253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
933 TGATGCTATTTATCAGTGCATGGCTGAGATAGCAAGGATCTATTTTATCTAGAGCCAG 992
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1254 CTTGCGGTAGTGGTGGCGAGGGGCTGCCAGCGCCCGGCTGCTGCTGCTGCTGCTGCTGCTG 1313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
993 ACTGACTGTAGTATGCTAGAGAGACAGACCCAGTCTCCCTATATGATGCTGTAAC 1052
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1314 GCTGAGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1053 CATGTCAAGCTCAGCCATCTTTTAGCTGGGAGGCGCACTTTATATTCAGACAAAGT 1112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1374 CATGGCTCTCTCTTCTACTACAAAGGAGGAGGAGTGGCAATGTGGAGTACCATGTT 1433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1113 CATTCCTATTTCTGTACACTACATGAAGAGCAAGGTTTAAATATATGAAGATATCAAGT 1172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1434 TGCAGTAACAAATGACACACAGAGCTGCAGTTTCGGGACCTGGAAACCCACACGATTA 1493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1173 AGTCATCGAAATGACAACTCATATATATATGATGCTTAGAGCTTGCAGCAATTA 1232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1494 TGAGTTTCTAGTGTGGCTCTATCCAGCTGGGGCCAGCCGAACTCCAGCCAGCCCT 1553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1233 TACTTTCTACATTTAGCATATATGCCAATGGAGCCAGCCAGATGCTCTGACCATGTGAC 1292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1554 GGTGATACACTGGAGAGATGCCCCAG 1580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1293 ACAGAATATCTAGAGGATGACCCAG 1319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 7

```
US-09-877-730-5
; Sequence 5, Application US/0987730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
```

```
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2715
; TYPE: DNA
; ORGANISM: homo sapiens
; US-09-877-730-5
```

Query Match 3.58; Score 217; DB 4; Length 2715;

Best Local Similarity 50.38; Pred. No. 6.1e-46;

Matches 647; Conservative 1; Mismatches 621; Indels 18; Gaps 4;

```
QY 300 AGCAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGCGAGCTATTTCCTG 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 279 ATACATCAGTGAGTGGAGGCGAGGAGGAGCAGTCCGATCGAAGATTTTATCATG 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 360 TCTGCCCCACAGCCCGCTAGAGTGTGGCCAGCCAGGTTGCTGTGTCAGGCTTGCAC 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 339 CTTGGCAATGAACAAATATGAGGCCATTCTTAGTCAAAAAGCTCATCTTGCTTTATCAAC 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 420 ACTCGAAGACTTCTCTCTGACCCCGAGTCCAGATTGTGGAGGAGAACGGGACGACG 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 399 TATTTCTGCAATTTGAAGTCCAGCCAAATTTCCACTGAGGTCACGAAGTGGAGTTGCTCG 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 480 CTTTGAATGCCACCAAGGCGCTTCACAGCCCGCATCAATTTCTGGGAAAGGACCAAGGT 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 459 ATTTCATGCAAGATTTTCATCCCGCTCTGTCAGTATAACATGCGAGTTCAATCGAC 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 540 GACCGTGTGAG---GAGCCCCGGCTCATCACTCTTCCCAGTGGCTCTCCAGATCCT 596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 519 AACTCTACCTATGACTATGAGACAGGATAAAGTCCCTACCAACAGAGATATTTCAGATCTA 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 597 AGATGTCAGGACAGTGTGAGGCTCTACCGCTGGTGGCCACCAATTCAGGCCGCCA 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 579 TGATGTGAGCCAAAGGAGTTCTGGAATTTATGTTGTTATGCTGCCACTGTAGGCCACCG 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 657 ACGATTCAGCCAGGAGCGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACCAG 716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 639 ACGTAAAGATGAGAGGCTCGCTAACTGTG-----ATTCAGCTAAGGAGTCAAAATC 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 717 GGGCAGGATGTGCTCAATTTGGCAGCCCGCAGAGAACACCGGTAGTGTCTGGACAGAA 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 693 CTTCCACACACCAACCAATTTATAGCAGGTCACAGAACATAACAACATCTCTTCATCAGAC 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 777 TGTAGTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 753 TGTAGTTTGGAAATGCAATGCGCCACAGGAAATCCCAAAACCAATCAATTTCTTGGAGCGGCT 812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 837 GGATGAAAGCTAT---CTCCAGGATGCTCATGCTTCTGGGCGCGGACCAATCTACTCAT 893
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 813 TGATCACAATCCATTTGATGCTTTAATCTCGGGTACTTGGAAATGCTATCTCATGAT 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 894 CGCCAGCGCGAGCTCGGCACTCTGGAGTCTATGCTGCGGAGCCAAACAGGCCCTCAC 953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 873 ATCTGATGTCAGGCTACAAACATGCTGGAGTATATGTTTGTGGGCCACTACCCCTGGCAC 932
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 954 GGTGACTTGGCACTGCGGCTGCTGAGTCTCGAGTGTGCTGCGCCAGCCAGCATCTCGCA 1013
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 933 ACGCAACTTTACAGTTGCTATGGCAACTTTAACTGATTAAGTCTCTCTCTTCAATTTGTA 992
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1014 GGCACCCGAGGCGCTCTCGCGAGCGCGGCGGAGCAACCGGCTGCTGCTGCTGCTGCTGCTG 1073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 993 ATGGCCAGAAAGTTTAAACAGGCTCTGAGCTGGCACTGCTCAATTTGTGTGAGGAGCA 1052
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1074 CGGGAGCCAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 1053 AGGAATCCCTCTCCCAAGATGTCAATGGTTGAAATAATGGAAGGAAGATACATTCGAATGG 1112
Qy 1134 GCGGTCAAGGTGCGAGGCGGTGGCGCAGCTTGTCTCATCTCAGATCGGCTCGCAGGA 1193
Db 1113 TAGAATTAATGATCAACAGT-----AAATGGTAAATTAACAGGATTAATTCCTGAAGA 1166
Qy 1194 CGTGGCTACTACAGGTGCGTAGCAGAAAACAGCGCGGAACTGCCCTGTGCCGCTGGCC 1253
Db 1167 TGATGCTATTATCAGTGTGATGCTGAGATAGCAAGGATCTATTTTATCTAGAGCCAG 1226
Qy 1254 CTTGGCGGTAGTGGCGAGGGGCTGCCAGGCGCCCGCAGCTCGGGTCAAGCCAGCC 1313
Db 1227 ACTGACTGTAGTGTGTCAGAAAGACAGACCCAGGTCTCCCTATAATGTATACATGCTGAAC 1286
Qy 1314 GCTGAGCAGCTCTCTGTGCTGTGGCTGGGAGCGGCTGAGTTGCACAGCGAGCAAT 1373
Db 1287 CATGTCAAGCTCAGCCATCTTTTAGCTGGGAGAGCCACTTTATTAATTCAGACAAGT 1346
Qy 1374 CATGGCTTCTCTCTCACTACCAAAAGCGAAGGGAGTGGACAATGTGGAGTACCAAGTT 1433
Db 1347 CATGGCTATTCTGTACACTACATGAAAGCAGAGGTTTAAATAATCAAGAGTATCAAGT 1406
Qy 1434 TGCAGTAAACATGACACACAGAGCTGCAGGTTTCGGGAGCTGGAAACCAACACGGATTA 1493
Db 1407 AGTCATCGGAATGACACAACTCATTTATTTATTTAGTACTTAGAGCTTGCAGCAATTA 1466
Qy 1494 TGAGTCTACGTGTGGCTACTCCAGCTGGGGGCGAGCGCACTCCAGCCAGCCCT 1553
Db 1467 TACTTTCTACATTTAGCATATATGCCAATGGGAGCCAGCCAGATGTCTGACCATGTGAC 1526
Qy 1554 GGTGCATACACTGACGATGTCCCCAG 1580
Db 1527 ACAGAACTCTAGAGGATGACCCAG 1553

RESULT 8

US-09-877-730-19

; Sequence 19, Application US/09877730

; Patent No. 6465632

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6465632a1 Human Phosphatases and Polynucleotides Encoding the

; FILE REFERENCE: LEX-0189-USA

; CURRENT APPLICATION NUMBER: US/09/877,730

; PRIOR FILING DATE: 2001-06-08

; PRIOR FILING DATE: 2000-06-07

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 19

; LENGTH: 2724

; TYPE: DNA

; ORGANISM: homo sapiens

US-09-877-730-19

Query Match 3.5%; Score 217; DB 4; Length 2724;

Best Local Similarity 50.3%; Pred. No. 6.1e-46;

Matches 647; Conservative 1; Mismatches 621; Indels 18; Gaps 4;

Qy 300 AGACAGCATGATGAGGAAGCTCTTAGGATCTGGAAGTCACTGAGGCGACCTATTCTGT 359

Db 45 ATACATCAGTGGTGGAGGCGAGGAGAGCAGTCCGATGAAGGATTTTATCAGTG 104

Qy 360 TCTGGCCACAGCCCGCTAGGAGTGGCGCAGCCAGGTTGCTGTGTTCAAGCTTGGCCAC 419

Db 105 CTTGGCAATGAACAAATATGGAGCCATTTCTTAGTCAAAAGGCTCATCTTGGCTTATCAAC 164

Qy 420 ACTCGAAGACTTCTCTCTGCAACCCCGAGTCCAGATTTGTGGAGGAGAAACGGGACACGACG 479
Db 165 TATTTCTGATTTGAAGTCCAGCAATTTCCACTCAGGTCCACGAAGGTGGAGTTGCTCG 224
Qy 480 CTTTGAATGCCACACCAAGGGCTTCCAGCCCCCATCATTAATCTTGGGAAAAGGACCAAGT 539
Db 225 ATTTGATCAAGATTTTCAATCCACCTCTCGAGTCATAACATGGAGATTCAATCGGAC 284
Qy 540 GACCGTGCCTGAG---GAGCCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCTCT 596
Db 285 AACTCTACCTATGACTATGCAAGGATAACTGCCCTACCAACAGGAGTATTGCAGATCTA 344
Qy 597 AGATGTCAGGACAGATGAGTGCAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCGCCA 656
Db 345 TGATGTCAAGCAAAAGGATTTCTGAAAATTTATGGTTGTTATGTGTCACATGTAGCCACCG 404
Qy 657 AGATTTCAAGCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTCGAGGCTACACG 716
Db 405 AGTAAAGATAGGAGGCTCGCTAACTGTG-----ATTCCAGCTAAAGAGTCAAAATC 458
Qy 717 GGGGCAAGGATGTGTCATTTGTGCGAGCCCCCAGAGAACACACGCTAGTGTCTGGACAGAA 776
Db 459 CTTCCACACACACCAATTTATAGAGGTCCACAGAACATAACAACTCTTTCATCAGAC 518
Qy 777 TGTAGTGTAGTGTGCTCTGCTGACCCCACTCTTTTGTGTCTCTGGGTCCGACA 836
Db 519 TGTAGTTTGTGAATGATGCGCACAGAAATCCCAACCAATCATTTCTTGGAGCCGCT 578
Qy 837 GGTGGAAGCCCTAT---CTCCACGATGTCTGCTGTGGCGGACCAATCTACTCAT 893
Db 579 TGTACAAATCCATTTGATGTCTTTAATACTCGGTAATTTGAAATGGTAAATCTCATGAT 638
Qy 894 CGCAGCGCGAGCTCGGCACTCTGGAGTCTATGTCTGCGGAGCAACAAAGCCCTCAC 953
Db 639 ATCTGATGTGAGCTACACATGCTGAGATATATGTTGTGGGCCACTACCCCTGGCAC 698
Qy 954 GCGTGAATTTGGCACTGCGGCTGTGAGCTCGAGTCTCGAGTGTGTCGCCAGCCATCTGCA 1013
Db 699 ACGCAACTTTACAGTTGCTATGGCAACTTTAACTGTATTAGTCTCTCTCATTTGTTGA 758
Qy 1014 GGCACCGAGGCGCTCTCGGAGCGGGCGGAGCACCGCGCTTCTGTTGCGCGGCGTC 1073
Db 759 ATGGCCAGAAAGTTTAAACAGGCTCGAGCTGCACTGCTCGATTTGTGTGTCAGGAGA 818
Qy 1074 CCGGAGCCACCGGCGCGCTGCACTGCGTGCACGACGAGATCCCGTTGCGACCCCAATGG 1133
Db 819 AGGAATCCCTCTCTCCAGATGTCATGTTGAAAATATGAAAGAGATATATTCGATGG 878
Qy 1134 GCGCGTCAAGGTGACGGCGGTGGCGGAGCTTTGGTCACTCATCTCAGATCGGCTGCAGGA 1193
Db 879 TAGAATTAATAATGTAACAACAGT-----AAATTTGGTAAATTAACAGATTTATCTGAAGA 932
Qy 1194 CGCTGGTACTACAGTGTGATGAGCAAAACAGCGGAGGAGTCCCTGTGCGCTGCGCC 1253
Db 933 TGATGCTATTATTTATCAGTGTGATGAGATAGCAAGGATCTATTTTATCTAGAGCCAG 992
Qy 1254 CTTGGCGGTAGTGTGCGGAGGCGCTGCCCGCGCGCTCGGCTCAGAGCCAGCCAGCC 1313
Db 993 ACTGACTGTAGTGTGAGAGACAGACCCAGTGTCTCCCTATATGATGATGCTGAAAC 1052
Qy 1314 GCTGAGCAGCTCTCTGTGCTGGCTGGGAGCGGCTGAGTGTGACAGCGAGCAAT 1373
Db 1053 CATGTCAAGCTCAGCCATTCTTTTAGCTGGAGAGGSCCACTTTATATTTACAGACAAGT 1112
Qy 1374 CATTTGGTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAAATGTGAGTATCCAGTT 1433
Db 1113 CATTTGCCCTATTCTGTACACTACATGAAGCAGAAAGGTTTAAATAATGAAGATATCAAGT 1172
Qy 1434 TGCAGTAAACAAATGACACACAGAGTGCAGGTTGGGAGCTGGAAACCCCAACCGGATTA 1493
Db 1173 AGTCATCGGAATGACACAACTCATTTATTTATTTGATGATCTTAGAGCTTCCAGCAATTA 1232
Qy 1494 TGAGTTCTAGTGTGGCTTACTCCAGCTGGGGGCGAGCGCAACCTCCAGCCAGCCCT 1553

[illegible]

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 21

; LENGTH: 2139

; TYPE: DNA

; ORGANISM: homo sapiens

US-09-877-730-21

Query Match 3.5%; Score 213.8; DB 4; Length 2139;

Best Local Similarity 47.0%; Pred. No. 3.6e-45;

Matches 793; Conservative 0; Mismatches 842; Indels 54; Gaps 2;

```
QY 1158 CGGACGCTGGTCAATCAGATCGGCTGCGAGGACGCTGCTACTACGAGTGGTAGC 1217
DB 60 CAGTAAATGGTAATTAACAGATATTCTGAAGATGATGCTATTATCAGTGCAATGCC 119
QY 1218 AGAAACAGGCGGGAAGTCTGTCGCGCTGCGCCCTGGCGGTAGTGGTGGCGAGG 1277
DB 120 TGAGAAATAGCCAAAGATCTATTTTATCTAGAGCCAGACTGATGTATGTGTCAGAAGA 179
QY 1278 GCTGCCAGGCGCCGACTCGGCTCACAGCCACGCGCTGAGCAGCTCCTCTGTGCTGGT 1337
DB 180 CAGACCCAGTCTCCCTATTAATGTACATGCTGAAACATGTCAAGCTCAGCCATTTCTTTT 239
QY 1338 GGCTGGAGCGGCTGAGTTGCACAGCGAGCAAAATCAATGGCTTCTCTTCACTACCA 1397
DB 240 AGCTGGGAGAGGCCACTTATAATTTCAGACAAAGTCATTTGCCCTATTCTGTACACTACAT 299
QY 1398 AAAGGCAAGGGAGTGGCAATGTGGAGTACAGTGTTCAGTAAACATGACACACAGAGA 1457
DB 300 GAAAGCAGAAAGGTTAAATTAATGAAGAGTATCAAGTAGTCAATCGGAAATGACACAACTCA 359
QY 1458 GCTGCAGTTTCGGGACCTGCAACCCACGAGTTATGAGTCTTACGTGGTGGCTACTC 1517
DB 360 TTATATATTATGATGACTTAGAGCTGCGACAAATTAATCTTCTACATTTAGCATATAT 419
QY 1518 CCAGCTGGGCGCCAGCGAAGCTCCAGCCGAGCCCTGGTGATACATCTGAGCAATGTCCC 1577
DB 420 GCCAATGGGAGCCAGCCAGATGCTGACCATGTGACACAGAACTCTAGAGGATGTTC 479
QY 1578 CAGGCGAGCAGCCAGCTTACCTTGTCCAGCCGACCCCTCGGACATCAGGGTGGGATG 1637
DB 480 CTTGAGACCTCTCGAAATTAAGTTTGAACAAGTCAAGTCCCACTGATATCTCATCTCTG 539
QY 1638 GCTGCCCTCCCTCCAGCTGAGCAATGACAGGTGCTCAAGTACAAAGTACAGTACGG 1697
DB 540 GCTGCCAATCCAGCCAAATATCGCGGGCCAGTGGTCTGTATGCTGTGCTTTCTCCG 599
QY 1698 TTTGGGAAGGAAGATCAGGTTTCTCCACCGAGGTGCTCGGAAATGAGACACAACTTAC 1757
DB 600 CCTAAGTACTGAGAAATTCATCCAAAGTTCTGGAGCTCCCGGGACCAACGATGATACCT 659
QY 1758 GTTAAATCTACTCAGCCAAACAAAGTGTACCGAGTCCGATTTTCAGCTGGGACTGGCGC 1817
DB 660 TTTGGGAAGGCTGAAACCTGACAGTGTCTACTCTGTTGGATTTACTGCTGCCACCAAGT 719
QY 1818 TGGCTATGAGTCCCTTCTCAGTGGATGACGACGACCACTGGTGTGACACACAGAG 1877
DB 720 GGGGCTGGGAGAGTCATGATATGGAATTCACATAGACGCGCCAAAGCTACAAGCGTGAA 779
QY 1878 CCATGTTTCCCTTGTCCCTTGAGAAATTTGAAGGTGAGGGCAAGATGAGTCCCTGGTGGT 1937
DB 780 AGCCCTTAAGTCTCCAGATTGCAATTTGGAGCTCTGAACTGTACCACTTTCTGTGAG 839
QY 1938 GTCATGGAGCGGCGCCCTCAACCCCAACAGATCTCTGGATACAAATCTTACTGGGAGA 1997
DB 840 GTGGCAGCAAGATGTAGAGGACACAGCTGTCTTTCAGGGGCTTACAGCTGTACTACAAGGA 899
QY 1998 GGTGGGAACAGAGGAGGAGGATGTGTACCGCGCCCGCCAGGGGCTGTGGAGATCAAGC 2057
DB 900 AGAAGGGCAGCAGAGAA----- 917
QY 2058 TTGGGAGCTCGGGCGCCCTGCGGCTGAAGAGAAAGTGAAGCAGTATGAATGACCCAGTT 2117
```

```
DB 918 -----TGGGCCCATTTTCTTGGATACCAAGGACCTTACTTATCTCTCAGTGGCTT 968
QY 2118 AGTCTCTGGCAGCGCTAGAGTGAAGCTCTGTAGCTTTTCAACAAACAGAGGACGGCTTA 2177
DB 969 AGACCCGAGAGAAATATATGTGAGACTCTCTGCTTACAAACAACATAGAGATGGCTTA 1028
QY 2178 CCGTCTGTGTGGAGGCAAGACGGAGAAAGCGCCAGCGCAGACCTGCTTATCCAGAG 2237
DB 1029 TCAGGCAGATCAGACTGTGAGCAGC---TCCAGGATGCGTGTCTGTTCGTGATCGCATGT 1085
QY 2238 GGGGCCACCGCTGCTCTGCGCCATGTCCAGCAGAGTCAACAGCTCACTTCCATTTG 2297
DB 1086 CCGTCTCTCCACACACCCACCATCTCTATGCGAAGGCTAAACCTCATCTTCCATCTT 1145
QY 2298 GCTTCGGTGGAGAAAGCCAGACTTTTACACTGTCAAGATTGTCAACTACACTGTAGCTT 2357
DB 1146 CTTGCACTGGAGAGGCTGCAATTCACCGCTGCAAAATCATTAATCAACCATCGCTG 1205
QY 2358 CGGCCCTCTGGGGCTCAGGAATGCTTCCCTGCTCACTACTATACAGCTCTGGAGAGA 2417
DB 1206 TAATCTGTGGCTGCGAGATGCTTTCTTGGTCTGTACTTCAAAACATCAGAACTCA 1265
QY 2418 CATTCATATTGGCGCTGAAACCATTTTACCAAGTACGAGTTTGGGTTAGCTTCCACGG 2477
DB 1266 CATGTTGGTTCAAGGCTTAGAAACCAAAACCAAAATACGAATTTGCGGTTGATATCATGT 1325
QY 2478 AGTGGATATGAGTGGCGCTTTGGCTCGCTGTAGAAAGCTCCACCTCGCTGACCGGCT 2537
DB 1326 GATCAGCTTTCAGTCTCTGGAGCCCTGTAGTCTACCAATCTACTCTTCCAGAGCAC 1385
QY 2538 TTCAACACCTCTCTTGACCTCGCTGAGCCCCCTGACACCATCCACCGTTGGTTACA 2597
DB 1386 AGCAGGCCCAACAGTTGGAGTAAAGTGACATTAATAGAGGATGACACTGCGCTGTTTC 1445
QY 2598 CTGGTGTCCCCCAAGGAGCCCAATGGTGAAGATTGTGGAGTATCTAATTTCTACAGCA 2657
DB 1446 TTGGAACCCCTGATGGCCAGAAACAGTTTGTGACCCGCTATCTATATGTCATC 1505
QY 2658 CAACACACCCAGCCGGAACACAGTGGACACTGCTCACACAGAGGGAACATCTTCAG 2717
DB 1506 TAGGAAGGCTGTGATTTGAGGAGTGGCAGGCTTACACCGTGAAGGGCAATAACCAT 1565
QY 2718 TGCAAGGTCCATGGCTAGAGAGTGACACTCGGTATTTTCTCAAGATGGGAGCCGAC 2777
DB 1566 GCGTTTGTAGAAACCTTGTAGCAGGAATGTGTACATTTGTCAAGATATCTGCAATCCA 1625
QY 2778 AGAGTGGGCGCTTGGCCCTTTTCCGCTTGGAGGATGTGATTAATCTCTGCAAGAGCAT 2837
DB 1626 TGAGGTGGGAGAGGAGCCCTTTTCAAAATTTCTGTGGAGCTGGCAGTACTTCCAAAGGAAC 1685
QY 2838 CTCAGACTC 2846
DB 1686 CTCGAAATC 1694
```

RESULT 11

US-09-877-730-27

; Sequence 27, Application US/09877730

; Patent No. 6465632

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding

; FILE REFERENCE: LEX-0189-USA

; CURRENT APPLICATION NUMBER: US/09/877,730

; PRIOR FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 60/210,607

; PRIOR FILING DATE: 2000-06-09

; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 2382
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-27

Query Match 3.5%; Score 213.8; DB 4; Length 2382;
Best Local Similarity 47.0%; Pred. No. 3.8e-45;
Matches 793; Conservative 0; Mismatches 842; Indels 54; Gaps 2;

```
QY 1158 CGGCAGCTGGTCATCACTCAGATCGGCTGCAGAGCGCTGGCTACTACCACTGGTACG 1217
DB 60 CAGTAAATGGTAAATTAACAGATTAATTCCTGAAGATGATGCTATTTATCACTGTCATGCG 119
QY 1218 AGAAACAGCGCGGAACTGCCTGTGCGCTGCGCCCTGGCGGTAGTGGTGGCGGAGG 1277
DB 120 TGAGATAGCCAAAGATCTATTTATCTAGAGCCAGATGACTGTAGTGTGTCAGAGA 179
QY 1278 GCTGCCAGCGCCCGACCTCGGGTCACAGCCACGCCGTGAGCAGCTCTCTGTGCTGGT 1337
DB 180 CAGACCCAGTCTCCCTAATGATGATGCTGAAACCATGTCAAGCTCAGGCCATTTCTTTT 239
QY 1338 GGCCTGGGAGCGGCTGGTGGTGCACAGCGAGCAATCATTTGGCTTCTCTCTCACTACA 1397
DB 240 AGCCTGGGAGGCGCCATTTAATTCAGACAAAGTCAATGGCTTATTTCTGTACACTACAT 299
QY 1398 AAAGCAAGGGAGTGGCAATGTGGAGTACCAGTTTCAGTAAACAATGACACCACAGA 1457
DB 300 GAAACAGAGGTTTAAATTAATGAAGAGTATCAAGTAGTCACTCGGAATGACAACTCA 359
QY 1458 GCTGAGGTTCGGGACCTGGAAACCAACACGGAATTAAGTTCAGTGGTGGGCTACTC 1517
DB 360 TTATATTATTGATGACTTAGAGCTGCGCAGCAATTAATTTCTTACATTTAGCATATAT 419
QY 1518 CCAGTGGGGGCGCGGACCTCCAGCCGAGCCCTGGTGGTCAATCACTGGAGATGTCCC 1577
DB 420 GCCAATGGGAGCGACCCAGATGTCTGACCATGTGACACAGAATACTCTAGAGGATGTTC 479
QY 1578 CAGCGCAGCAGCCAGCTTACTTGTCCAGCCCAACCCCTCGGACATCAGGGTGGCATG 1637
DB 480 CTGAGACTCTCTGAAATAGTTTGAAGTGGAAAGTCCCACTGATATTTCTCATCTCCG 539
QY 1638 GCTGCCCTGCGCTCCAGCTGAGCAATGGACAGGTGCTGAAGTACAAGATAGATACGG 1697
DB 540 GCTGCCAATCCAGCCAAATATCGCGGGGCCAAGTGGTGTGTATCGCTTCTTCCG 599
QY 1698 TTGGGGAGGAAGATCAGGTTTCTCCACGAGGTGCTTGAATAGACACAACTTAC 1757
DB 600 CTAAGTACTGAGAAATCAATCCAAAGTTCTGGAGCTCCCGGGACCAAGCATGAGTACCT 659
QY 1758 GTTAAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGGATTTTCAGCTGGCACTGGCGC 1817
DB 660 TTGGAAAGCCCTGAAACCTGACAGTGTCTACCTGGTTCCGATTTACTGTGCCACAGAGT 719
QY 1818 TGGCTATGAGTCCCTTCTCAGTGTGATGCAGCAGGACACTGGTGTGCACAAACAGAG 1877
DB 720 GGGGCTGGAGGATCATCAGTATGAGCTTTCATAGGAGCGCCAAAGCTACAGGCTGGA 779
QY 1878 CCATGTTCCCTTTGCCCCCTGCGAATTTGAAGGTGAGGCAAGATGAGAGTCCCTGGTGGT 1937
DB 780 AGCCCTAAGTCTCCAGAGTTGCAATTTGGAGCCTCTGAACCTGTACCAACCATTTCTGTGAG 839
QY 1938 GTCATGGCAGCGGCCCCCTCAGCCCAACCAAGATCTCTGGATACAACTCTACTGGGGAGA 1997
DB 840 GTGGCAGCAAGATGTAGAGGACACAGCTGCTATTTCAGGGCTACAGCTGTACTACAAGGA 899
QY 1998 GTTGGAAACAGAGGAGGAGGAGGAGTGGTGAACCGCCCCCAGGGGGTCTGTGGAGATCAAGC 2057
DB 900 AAGGGGAGCAGGAGAA----- 917
QY 2058 TTGGAGCTCGGGGCCCCGTGGGCTGAAGAGAAAGTGAAGCATGATGAACCTGACCCAGTT 2117
```

```
DB 918 -----TGGGCCCAATTTTCTTGGATACCAAGGACCTACTCTATACTCTCAGTGGCTT 968
QY 2118 AGTCCCTGGCAGGCGCTACGAGGTGAAGCTCGTAGCTTTTCAACAAACACAGGAGCGGCTA 2177
DB 969 AGACCCAGAGAAATATCATGTGAGACTCTCTGGCTTTACAAACATAGACGATGGCTA 1028
QY 2178 CGCTGCTGTGTGGAAGGCAAGACGGAGAGGCGCCCAACGACAGCTGCTCCATCCAGAG 2237
DB 1029 TCAGGCAGATCAGACTGTCTCAGCAC---TCCAGGATGGGTGTCTGTCTGTGATCGCATGGT 1085
QY 2238 GGGGCCACCGCTGCTCTCGCCCATGTCCAGAGTGCACGAGGTCAACAGCTCCACTTCCATTG 2297
DB 1086 CCCTCTCCACACACACCCCATCTCTATGGAAGGCTTAACACCTCATCTTCCATCTT 1145
QY 2298 GCTTTCGGTGAAGGAAGCAGACTTTTACCACTGTCAAGATTGTCAACTACACTGTACGCTT 2357
DB 1146 CCTGCATCTGGAGGAGGCTGCATTCACCGCTGCACAAATCATTAACCTACACCATCCGCTG 1205
QY 2358 CGGCCCTCTGGGGCTCAGGAATGCTTCCCTGGTCACTCTATACCAGCTCTGGAGAGA 2417
DB 1206 TAATCTCTGTGGCTTGCAGAAATGCTTCTTGGTCTGTGTACCTTCAAAACATCAGAACTCA 1265
QY 2418 CATCTCTCATTTGGCGGCTGAAACCAATTTTACCAAGTACGAGTTTGGGTACAGTCCACGG 2477
DB 1266 CATGTTGGTTCGAAGTCTAGAACCAAAACAAATAAGAAATTTGGCCGTTCGATTTACATGT 1325
QY 2478 AGTGGATATGATGGGCGCTTTGGCTCCGTGTAAGACGCTCCACCTGCTGACCCGCGCC 2537
DB 1326 GGATCAGCTTTCAGTCTCTGGAGCCCTGTAGTCTACATTTCTACTCTTCCAGAGACCC 1385
QY 2538 TTCAACACCTCTCTTCAAGCTGCGCTGAGCCCTGAGCCCTGACCACTCCACCGTTCCGTTTACA 2597
DB 1386 AGCAGGCCACCACTGGTGAAGTAAAGTGAATTAATAGAGGATGACACTGCGCTGGTTTC 1445
QY 2598 CTGCTGTCCCTCCAGGAGCCCAATGGTGAAGATTTGTGAGATATCTAAATCTCTACAGCA 2657
DB 1446 TTGAAACCCCTCTGATGGCCAGAAACAGTTGTGACCCGCTATATCTATATATGCAATC 1505
QY 2658 CAACCAACACCCAGCCCAACACCACTGCACTGCTCACCACAGAGGAAACATCTTCAG 2717
DB 1506 TAGGAAGGCTGGATTTGAGAGAGTGGCAGGTCTTACACCGTGAAGGGCAATTAACCAT 1565
QY 2718 TGCAGAGTCCATGCGCTTAGAGGTGACACTCGGTATTTCTTCAAGATGGAGCGCCGAC 2777
DB 1566 GCTTTGTAGAAAACTTGGTAGCAGGAATATGTCTACATTTGTCAAGATATCTGCATCCAA 1625
QY 2778 AGAGTGGGGCTGGGGCTTTTCCCGCTTGCAGGATGTGATTAATCTCTGCAAGAGACATT 2837
DB 1626 TGAGTGGGAGAGGAGCCCTTTTCAAAATTTCTGTGGAGCTGGCAGTACTTCCAAAGGAAC 1685
QY 2838 CTCAGACTC 2846
DB 1686 CTCTGAATC 1694
```

RESULT 12

US-09-877-730-23

; Sequence 23, Application US/09877730

; Patent No. 6465632

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding :

; FILE REFERENCE: LEX-0189-USA

; CURRENT APPLICATION NUMBER: US/09/877,730

; CURRENT FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 60/210,607

Db 1113 TAGAATTAAATGTA 1127
RESULT 15
US-09-877-730-25
; Sequence 25, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6456632el Human Phosphatases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-25

Query Match 1.9%; Score 117; DB 4; Length 1644;
Best Local Similarity 48.4%; Pred. No. 3.6e-20;
Matches 355; Conservative 0; Mismatches 375; Indels 3; Gaps 1;

QY 2114 AGTTAGTCCCTGGCAGGCGCTAGAGGTGAAGTCTGTAGCTTTCAACAAACACGAGGAG 2173
Db 470 AGGATGACCCAGAGAAATATCATGTGAGACTCTCTGGCTTACAAACATAGACGATG 529
QY 2174 GCTACGCTGTGTGTGGAGGGCAAGACGGAGAGCGCCACGCCAGACCTGCCTATCC 2233
Db 530 GCTATCAGGAGATCAGATCTGTCAGAC---TCAGAGTGGTGTCTGTTGTGATCGCA 586
QY 2234 AGAGGGGGCCACCGCTCTCTGCCCATGTCCACGAGAGTCAAAACAGCTCCACTTCCA 2293
Db 587 TGGTCCCTCTCCACCAACCCACCATCTCTATGGAAGGCTAACACCTCATCTTCCA 646
QY 2294 TTTGGCTTGGTGGAGAGACCGACATTTACCACTGTCAAGATTGTCAACTACACTGTAC 2353
Db 647 TCTTCTGCACTGGAGAGGCGCTGCATTTACCGCTGCACAAATCATTAACCTACACCATCC 706
QY 2354 GCTTGGCCCTGGGGGCTCAGGAATGCTTCCCTGGTCACTACTATACAGCTCTGGAG 2413
Db 707 GCTGTAATCTGTGGCTGCAGAAATGCTTTGTTGTTCTGTACCTTCAAAACATCAGAA 766
QY 2414 AAGACATTTCTCATTTGGCGGCTGAAACATTTACCAAGTACGAGTTTGGGTACAGTCCC 2473
Db 767 CTCACATTTGGTTCAAGGTCTAGAACCAACCAAAATACGAATTTGCGTTTGGATTAC 826
QY 2474 ACGGAGTGATATGAGATGGGCCCTTTGGGCTCCGTGTGTAGAAACGCTCCACCCCTGACC 2533
Db 827 ATGTGGATCAGCTTTTCAGTCTTGGAGCGCTGTAGTCTACCATTTCTACTTCTCCAGAAG 886
QY 2534 GGCCTTTCAACACCTCTCTGACCTGGCTGAGCCCTCGACCCATCCACCGTTGGT 2593
Db 887 CACGAGAGGCCACCAAGTTGAGTAAAGTGACATTAATAGAGATGACACTGGCCCTGG 946
QY 2594 TACACTGTGTGTCCCGCCACCGAGCCCAATGGTGTGAGATTGTGGAGTATCTAAATTTCTTACA 2653
Db 947 TTTCTTTGGAACCCCTGTATGGCCAGAAACAGTTGTGACCCGCTATCTATCTATATG 1006
QY 2654 GCAACACCAACACCCAGCCCGAAACACGATGGACACTGTCTCAACACAGAGGGAAACATCT 2713
Db 1007 CATCTAGGAAGGCGCTGTGATTTGAGGAGATGGCAGGTCTTTACACCGTGAAGGGCAATAA 1066

Search completed: October 15, 2003, 04:46:31
Job time : 383.754 secs

QY 2714 TCAGTGCAGAGGTCCATGGCTTAGAGAGTGACACTCGGTATTTCTTCAAGATGGAGCCC 2773
Db 1067 CCATGGCTTTGCTTAGAABACTTTGGTAGCAGGAAATGTGTACATTGTCAAGATATCTGCAT 1126
QY 2774 GCACAGAGGTGGGGCTTGGGCCCTTTTCCCGCTTGCAGGATGTGATTTACTCTGCAAGAGA 2833
Db 1127 CCAATGAGGTGGGAGAGGAGCCCTTTTCAAAATTTCTGTGGAGCTGGCAGTACTTTCCAAGG 1186
QY 2834 CATTTCTCAGACTC 2846
Db 1187 AAACCTCTGAATC 1199

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 14, 2003, 06:11:05 ; Search time 7190 Seconds
(without alignments)
7123.623 Million cell updates/sec

Title: US-09-754-997A-2

Perfect score: 6604

Sequence: 1 MARADTCRGLLVLTFCLLSA.....CPLTVSPSLPAPVSSAQP 1252

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09754997/runat_14102003_061100_9706/app_query.fasta_1.1415
-DB=GenEmbl -QFWT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOFCL=0 -LOPEXT=0
-UNITS=dbs -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09754997@cgn_1_1_4823@runat_14102003_061100_9706 -NCPU=6 -ICPU=3
-NO MAP -LARGESUBQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DSLOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	6604	100.0	3756	6	AX191303	Sequence
2	6604	100.0	6176	6	AX191259	Sequence
3	6604	100.0	6176	10	AF176694	Mus muscu
4	6532	98.9	6219	10	AB052620	Mus muscu
5	6521.5	98.8	6222	10	AB052621	Mus muscu
6	6351.5	96.2	6301	10	AK122535	Mus muscu
7	5746	87.0	6485	9	AB052622	Homo sapi
8	5610	84.9	3741	6	AX418406	Sequence
9	4903	74.2	2796	6	AX191261	Sequence
10	4361.5	66.0	5549	9	AB046848	Homo sapi
11	3125.5	47.3	158055	2	AC110235	Mus muscu
12	3064.5	46.4	224227	2	AC106509	Rattus no
13	2557	38.7	176282	2	AC068507	Homo sapi
14	2557	38.7	198295	9	AC105129	Homo sapi
15	2466.5	37.3	115580	10	AC112162	Mus muscu
16	2321	35.1	115135	10	AC112161	Mus muscu
17	1747.5	26.5	3453	6	AR237556	Sequence
18	1746.5	26.4	3210	6	AR237553	Sequence
19	1746.5	26.4	3874	6	AR237568	Sequence
20	1691.5	25.6	2976	6	AR237558	Sequence
21	1691.5	25.6	3219	6	AR237561	Sequence
22	1472	22.3	825	6	AX191263	Sequence
23	1384	21.0	2715	6	AR237555	Sequence
24	1384	21.0	2958	6	AR237557	Sequence
25	1329	20.1	2481	6	AR237560	Sequence
26	1329	20.1	2724	6	AR237562	Sequence
27	1144	17.3	2139	6	AR237563	Sequence
28	1144	17.3	2382	6	AR237566	Sequence
29	1070.5	16.2	2967	9	BC042054	Homo sapi
30	1067	16.2	2700	6	AX478857	Sequence
31	1046	15.8	3198	10	BC053057	Mus muscu
32	1031.5	15.6	3146	10	AF026465	Mus muscu
33	978	14.8	4784	5	GU07644	Gallus gall
34	953.5	14.4	4631	9	HSU72391	Human neog
35	951.5	14.4	5297	6	AX587800	Sequence
36	951.5	14.4	5297	9	HSU61262	Human neog
37	947	14.3	4641	10	RNU68726	Rattus norv
38	945.5	14.3	5414	10	MMNEOGEN	M.musculus
39	943	14.3	4344	6	AX054981	Sequence
40	943	14.3	4608	6	AR153583	Sequence
41	943	14.3	4608	9	HSDCCG	H.sapiens D
42	940.5	14.2	4517	10	MMDCCTSG	M.musculus
43	932.5	14.1	4846	5	XLU10986	Xenopus lae
44	931	14.1	4924	10	RNU68725	Rattus norv
45	921.5	14.0	2729	9	AK095529	Homo sapi

ALIGNMENTS

RESULT 1

AX191303
 LOCUS AX191303 3756 bp DNA linear PAT 15-AUG-2001
 DEFINITION Sequence 45 from Patent WO0149714.
 ACCESSION AX191303
 VERSION AX191303.1 GI:15209554
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Salbaum, J. M.
 Nope polypeptides, encoding nucleic acids and methods of use
 TITLE Patent: WO 0149714-A 45 12-JUL-2001;
 JOURNAL Neurosciences Research Foundation Inc. (US)
 FEATURES
 source 1..3756
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 BASE COUNT 774 a 1167 c 1105 g 710 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.74e-260 Length: 3756
 Score: 6604.00 Matches: 1252
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-754-997A-2 (1-1252) x AX191303 (1-3756)
 QY 1 MetAlaArgAlaAspThrGlyArgGlyLeuLeuValLeuThrPheCysLeuLeuSerAla 20
 DB 1 ATGGCGGGCGGACACGGCGCGGCTCTGGTGTCTGACCTTCTGCGCTGCTCGCG 60
 QY 21 ArgGlyGluLeuProLeuProGlnGluThrValIleLeuSerCysAspGluGlyPro 40
 DB 61 CGCGGGAGCTGCATTTGCCAGGAGACACTGTCTAGCTGAGCTGTGATGAGGACCC 120
 QY 41 LeuGlnValIleLeuGlyProGluGlnAlaValIleLeuAspCysThrLeuGlyAlaThr 60
 DB 121 CTGCAAGTGTATCTGGGCGCTGAGCAGGCTGTGTCTGAGTGCATCTTTGGGGGTACA 180
 QY 61 AlaAlaGlyProProThrArgValThrTrpSerIleAspGlyAspThrValLeuGluHis 80
 DB 181 GCTGTGGGCTCCGACCAAGGTGACATGGAGCAAGGATGGAGACACTGTACTAGAGCAT 240
 QY 81 GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu 100
 DB 241 GAGAACCTGCACCTGTCTACCAATGGCTCCCTGTGGCTGTCTACCCCTAGAGCAAGAA 300
 QY 101 AspSerAspAspGluGluAlaLeuArgIleTrpIleValThrGluGlySerTrpSerCys 120
 DB 301 GACAGCGATGATGAGGAGCTCTTAGATCTCGAAGGTCACTCAGGGGAGCTATTCTGT 360
 QY 121 LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValIleValLeuAlaThr 140
 DB 361 CTGGCCACACAGCCGCTAGGAGTGTGGCCAGCAGGTTCTGTGTCTCAAGCTTGGCCACA 420
 QY 141 LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg 160
 DB 421 CTGAGAGACTTCTCTGCAACCCGAGTCCAGATTGTGGAGAGAACGGGACAGCAGCG 480
 QY 161 PheGluCysHisThrIleGlyLeuProAlaProIleIleThrTrpGluIleAspGlnVal 180
 DB 481 TTGTGATGCCACACCAAGGCGCTTCCAGCCCTCATCTACTTGTGGAAAGGACAGGTG 540
 QY 181 ThrValProGluGluProArgLeuIleThrLeuProIleTrpLeuLeuGlnIleLeuAsp 200
 DB 541 ACCGTGCTGAGGAGCGCGCTCATCACTTCTCCCAAGTGGCTCTCTCCAGATCTTAGAT 600

QY 201 ValGlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArg 220
 DB 601 GTCCAGGACAGTGTACAGGCTCTTACCGTGTGGCCACCAATTACGCCGCCCAACA 660
 QY 221 PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly 240
 DB 661 TTCCAGCCAGGAGGCTCTGCTCACTGTGGCCCTCAGAGGCTCTTTGGAGCTTACCAAGGG 720
 QY 241 GlnAspValIleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnVal 260
 DB 721 CAGGATGTGTCTTTGTGGCAGCCCGAGAGAACACACGGTAGTGTCTGGACAGATGTA 780
 QY 261 ValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp 280
 DB 781 GTGATGAGTGTGGTCTCTGCTGACCCCACTTGTGTCTCTGGGTCCGACAGGAT 840
 QY 281 GlyIleProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuLeuAsp 300
 DB 841 GGAAGAGCTTATCTCCAGGATGTCTGCTTGTGGGCCCGGACCAATCTACTATCCGACG 900
 QY 301 AlaGlnProArgHisSerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAsp 320
 DB 901 GCGGAGCTCGGCACTCTGGAGTCTATGTCTCCGAGCGCACAGCCCTCTACGGTGAC 960
 QY 321 PheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro 340
 DB 961 TTTCGCCACTGCGGCTGTGAGCTCCGAGTGTCTGCTGCCCGCAGCATCTCCGAGGACCC 1020
 QY 341 GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu 360
 DB 1021 GAGGCGCTCTCGGAGCGCGGCGGACCGCGCGCTTGTGTGGCGGCGTCCGCGGAG 1080
 QY 361 ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgVal 380
 DB 1081 CCAGCGCCCGCTGTCACTGGCTGCAGAGCGAGATCCCTGCGACCCCAATGAGCGCGCTC 1140
 QY 381 LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly 400
 DB 1141 AAGGTGAGGCGGCTGGCGGAGCTTGTCTCATCTCAGTCCGCTCGCGCTCGAGACGCTGGC 1200
 QY 401 TyrTrpGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAla 420
 DB 1201 TACTACAGTGTGTAGCAGAAACAGCGGGAACTGCTGTGCGCTGTGCGCTCGCGCTGGCG 1260
 QY 421 ValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer 440
 DB 1261 GTAGTGTGCGGAGGCGCTGCCAGCGCCCTGACTCGGCTCAGCCAGCGCTGGC 1320
 QY 441 SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGly 460
 DB 1321 AGCTCTCTGTCTGTGGCTGGGAGCGGCTGAGTTGTCACAGCGAGCAATCATTTGGC 1380
 QY 461 PheSerLeuHisTrpGlnLysAlaArgGlyValAspAsnValGluTrpGlnPheAlaVal 480
 DB 1381 TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAAATGTGGAGTACCAAGTTTGCAGTA 1440
 QY 481 AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPhe 500
 DB 1441 AACAAATGACACACAGAGCTGCGAGTTCGGGACCTGGAACCCCAACACCGATTATGATTC 1500
 QY 501 TyrValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis 520
 DB 1501 TAGTGTGGCTTACTCCAGCTGGGGGCGAGCGGACCTCCAGCCCGCCTGTGGTGCAT 1560
 QY 521 ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProSer 540
 DB 1561 ACACTGACGATGTCCCGAGCGCAGCACTTACCTGTGTCTGAGCCCAACCCCTCG 1620
 QY 541 AspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLys 560
 DB 1621 GACATCAGGCTGGCATGGCTGGCCCTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAG 1680
 QY 561 TyrLysIleGluTrpGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly 580

Db 1681 TACAGATAGAGTACGGTTTGGGGAAGGAAGATCAGGTTTCTCCACCGAGGTCCCTGGA 1740
Qy 581 AenGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgile 600
Db 1741 AATGAGACAACTTACGTTAACTCACTTCAGCCCAACAAGGTGTACCGAGTCCGAT 1800
Qy 601 SerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrPro 620
Db 1801 TCAGTGGCACTGGCGCTGGCTATGGAGTCCCTTCTCAGTGGATGCAAGCAAGGACACCT 1860
Qy 621 GlyValHisGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys 640
Db 1861 GGTGTGCACAAACAGGACATGTTCCCTTTGCCCTGAGAAATGAAGGTGAGGCAAG 1920
Qy 641 MetGluSerLeuValSerTrpGlnProProHisProThrGlnLysSerGlyTyr 660
Db 1921 ATGGAGTCCCTGGTGGTGTATGGACGCGCCCTCACCCACCCAGATCTCTGGATAC 1980
Qy 661 LysLeuTyrTrpGlyGluValGlyThrGluGluAlaAseGlyAspArgProProGly 680
Db 1981 AAACCTACTGGGAGAGGTGGGAACAGAGGAGGAGGAGATGGTGCACCGCCCGCAGGG 2040
Qy 681 GlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGln 700
Db 2041 GGTGTGGAGATCAAGCTTGGGACGTGGGCGCCCTGGCGCTGGAAGAAAGTGAAGCG 2100
Qy 701 TyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsn 720
Db 2101 TATGACTGACCCAGTATGCTTCCCTGGCAGCGCTACGAGGTGAAGCTCTGAGCTTTCAAC 2160
Qy 721 LysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrPro 740
Db 2161 AAACACAGAGGACGCTACGCTGCTGTGTGAAGGCGCAAGACGAGAGGCGCCACGCCA 2220
Qy 741 AspLeuProLysGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsn 760
Db 2221 GACCTGGCTATTCAGAGGGGCGCCACCGCTGGCTCTCTGCCCATGTCCACGAGAGTCAAC 2280
Qy 761 SerSerThrSerLeuTrpLeuArgTrpLysLysProAspPheThrThrValLysVal 780
Db 2281 AGCTCCACTTCATTTGGCTTGGTGGGAAGGCGAGCTTTACACTGTCTCAAGATTGTC 2340
Qy 781 AsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyr 800
Db 2341 AACTACACTGTACGCTTGGCCCTCGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT 2400
Qy 801 ThrSerSerGlyGluAspLysLeuLysGlyLysLeuLysProPheThrLysTyrGluPhe 820
Db 2401 ACCAGCTCTGGAGAAGACATTTCTATTGGCGGCTGAAACCATTTTACCAAGTACGAGTTT 2460
Qy 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840
Db 2461 GCGGTACAGTCCACGAGTGGATATGGATGGGCGCTTGGCTCGCTGTAGAACGCTCC 2520
Qy 841 ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
Db 2521 ACCCTGCTGACCGGCTTCAACACTCTTCTGACCTGGCGCTGAGCGCCCTGACACCA 2580
Qy 861 SerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluLysValGluTyr 880
Db 2581 TCCACCGTTTGGTTACATGTTGTGTCCTCCCGACGAGGCCAATGGTGGATTTGTGGATAT 2640
Qy 881 LeuLeuLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThr 900
Db 2641 CTAAATTTCTACAGCAACCAACACACCGGCGCCGGAACACCAAGTGGACATCTCACCACA 2700
Qy 901 GluGlyAsnLysPheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
Db 2701 GAGGGAACATCTTCAGTTCAGAGGTTCATGGCCCTAGAGAGTGCACCTCGGTATTTCTTC 2760
Qy 921 LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValile 940

Db 2761 AAGATGGAGCCCGCACAGAGGTGGGCGCTTGGGCGCTTTTTCGCTTGCAGGATGTGATT 2820
Qy 941 ThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyLysValileVal 960
Db 2821 ACTCTGCAAGAGACATTTCTCAGACTCTCTGGATGTGCACGCGCTCGACGGGCATCATCTGT 2880
Qy 961 GlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
Db 2881 GGTGTCTGGCTGGGCGCTTCTCTGCTCTCTGGCCCTGCATGTGTCTGTGCGCTACGACAAAGC 2940
Qy 981 SerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeu 1000
Db 2941 TCCACACAGGAGCGCTTCCCGATTTGTCTCTCAGGCAACCCAGGAAACCCAGCGCTC 3000
Qy 1001 TyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeu 1020
Db 3001 TACACAAGAGCTCGCTTGGGCTCTCCAGTGTCTCTGCTGCCATGATGTTGGAGTCCCTC 3060
Qy 1021 ValHisProArgProGlnAspTrpSerProProSerAspValGluAspLysAlaGlu 1040
Db 3061 GTGCATCTCTGCTCCCGAGGATTTGGTCCCAACCCCTCAGATGTGGGAACAAGGCTGAA 3120
Qy 1041 ValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysValile 1060
Db 3121 GTACACAGCTTATGGTGGCAGTGTCTCAGATTGCCGGGCGCACTCCAAGAGAAAGATC 3180
Qy 1061 SerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuPro 1080
Db 3181 TCCTGGGCTCAGGCAGGCGGACCAACTGGGCGAGGCTCTGGGCGAGCTGTGAGTCCCTC 3240
Qy 1081 GlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
Db 3241 CAGGATAGTGTCCAAAGCGCGCTCTGACCGCTCTGCTGCTCAGCGGGAACCGGG 3300
Qy 1101 GlnThrLeuLeuLeuGlnAlaLeuValTyrAspGlyLysLysSerAsnGlyArgLysLys 1120
Db 3301 CAGACACTGCTGTCTGCAAGCGCTGTGTATGACGGCATTAAGAGCAACGGGAGAAAGAG 3360
Qy 1121 ProSerProAlaCysArgAsnGlnValGluAlaGluValLysValHisSerAspPheGly 1140
Db 3361 CCGTCCCGCGCTCGAGGAATCAGGTGGAAGTGAAGTCAATGTCCTCCTCGACTTCGGT 3420
Qy 1141 AlaSerLysGlyCysProAspLeuHisGlnAspLeuGluProGluGluProLeuThr 1160
Db 3421 GCATCCAAAGATGTCTCTGACCTCCACTCCAGACCTGGAGCCAGAGGAACCACTGACT 3480
Qy 1161 AlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeu 1180
Db 3481 GCAGAGACTCTGGCTTCCACGCTCTGGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTG 3540
Qy 1181 GlyArgGluLeuGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
Db 3541 GGCAGGAGCTGGGAGGTGTCACCAACCAACCGTGGGCGCAGAGAGGCTCACCTGCTTG 3600
Qy 1201 ProGluAlaAlaSerAlaSerCysSerCysAspLeuGlnProSerThrAlaLysGlu 1220
Db 3601 CCAGAGACGACGAGTGGCTCTCTGCTCTGCTCTCAGACCTCCAGCCACGACTGTATAGAG 3660
Qy 1221 GluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSer 1240
Db 3661 GAGGCGCTCTGGAAAGCTCCAGCCCAAGGCGCTGTGTCTCTTAAAGTCAAGTCAAGCCCAAGC 3720
Qy 1241 LeuProArgAlaProValSerSerAlaGlnValPro 1252
Db 3721 CTTCCAGGCGCTCTGTCTCTCTCTGCTCAGGTCCCTC 3756

RESULT 2

AX191259
LOCUS AX191259
DEFINITION Sequence 1 from Patent W00149714.
ACCESSION AX191259
VERSION AX191259.1 GI:15209510
KEYWORDS

AX191259 6176 bp DNA linear PAT 15-AUG-2001

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS	Salbaum, J.M.
TITLE	Nope polypeptides, encoding nucleic acids and methods of use
JOURNAL	Patent: WO 0149714-A 1 12-JUL-2001;
FEATURES	Neurosciences Research Foundation Inc. (US)
source	Location/Qualifiers
	1..6176
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
CDS	1..3759
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC51120.1"
	/db_xref="GI:15209511"
	/translation="MARADPGRLLLVTFCLLSARGELPLPQETTVKLSCDEGPGLOVYLGFEQAVLVKDTLGATAAGPTVTVKSGDVTLEHENLHLLPNSLMLSSPLEQSDSDEEARLVKVTGSLAHSPGLVWSQVAVVKLATLDFLSPESQIVBENGSTARFECHTKGLPAPIITWEKDVYTPPEERPLITLPKWLLQILDVQSDSDAGSYRCVATNSARGQSEASLTVALRGLSEALTEFGQDVIVAAPEVTTVSGGNVMBVCASADPTPFVSWVRQDKPISDVIVLGRTNLLIASAQHSGVTVCRANKPLTRDFATAABLRVLAAPAIASQPEALSRCTASARFVCRASGEPRLALHWDGIPLRNGRVKVGQGGGSLVITQIGLOAGYTCQVAENSAGTACAAAPLAVVVBREGLPSAPTRVTATPLSSSVLVAWERPELHSEQIIGFSLHYOKARGVDNVVEQFAVNNDTTELQVRDLBPNTDYBFFVVAYSOLGASSTSPALVHTLDDVPSAAPOLTLSSPNPSDIRVAMLPPLSSLNGOVLVKYIEYGLKREDQVFSTEFGNETQTLNLSLPKNKYRVRIISAGTAGYGVPGQMQRHTPGVHNQSHVFPAPAEILKVRKMSLVSWSQPPHPPTQISGYKLWYGEVTEBBAQDGRPPGGRQDAQWDVRLKKKQYKQYELTQLVQPPRPYEVKLAFNPHGEDYAAVMKGKTSKATPDLPTQRGPPPLPAHFVHAESNSSTIWLKMKKDPFTTVKIYNYTVRFGPMLGRNASLVTYTSQGBDILIGGLKPTKYEFVQSHGVMDGPFSGSVVERSTLPDRPSTPPSDLRSLPLTPTSVLWKCPTPEBNGIEVLLYSNNHTQPEHQTWLTTEGNIPSAEBVHGLS DTRYFPFKGARTVEGPFPSRLQDVITLQETPDSLDLVHAVTGIIVGVCLGLCLLACMCAGLRQSSHREALPGSSSLPGNPLAYTRALRPPSPVPAHLESLHPRCDWSPPLPSDVEDKAEVHSLMGSGSDCHGRXKRI SWAQAGGNWAGSWAGCGLPQSGSGPRPALTRALLPACTGOTLLIQAALVYDGIKSNGRKPSPACRNOVZAEVIVHSDFFGASKGCPDLHLDQLEPPEPLTAETLPTSGAVDLSDQADWLGRGLGCGQPTTSPGRLTCLPBAASLSCSSDLPQSTAIERAPGKSCQPKALCLTVVPSPLPRAPVSSAQVP"
BASE COUNT	1316 a 1756 c 1634 g 1470 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3,04e-260 Length: 6176
Score:	6604.00 Matches: 1252
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	6 Gaps: 0
US-09-754-997A-2 (1-1252) x AX191259 (1-6176)	
Qy	1 MetAlaArgAlaAepThrGlyArgGlyLeuLeuValLeuThrPheCysLeuLeuSerAla 20
Db	1 ATGCGCGGGCGGACACGGGCGCGGGCTCTGCTGCTGCACCTTCTGCCTGCTGCCG 60
Qy	21 ArgGlyGluLeuProLeuProGlnGluThrThrVallysLeuSerCysAspGluGlyPro 40
Db	61 CGCGGGGAGCTGCCATTGTGCCCCAGGAGACAACCTGTCAGCTGAGCTGTGATGAGGGACCC 120
Qy	41 LeuGlnValIleLeuGlyProGlnGlnAlaValValLeuLeuAspCysThrLeuGlyAlaThr 60
Db	121 CTGCAAGTGATCTCTGGGCGCTTGACAGCGCTGTGTGCTGTGACTGCATCTTTGGGGGCTTACA 180
Qy	61 AlalaGlyProProThrArgValThrTrpSerLyAspGlyAspThrValLeuGluHis 80
Db	181 GCTGCTGGGCTCCGACAGGGTGCATGGAGCAAGATGGAGACACTGTACTAGAGCAT 240
Qy	81 GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu 100
Db	241 GAGAACCTGCACCTGTACCCAAATGGCTCCCTGTGGCTGTCTCACCCTCCTAGACAGAA 300

QY 461 PheSerLeuHisTyrGlnIysAlaArgGlyValAspAsnValGluTyrGlnPheAlaVal 480
DB 1381 TTCTCTCTTCACTACCAAAAGCAAGGGAGTGGCAATGTGGAGTACAGATTGACGTA 1440
QY 481 AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPhe 500
DB 1441 AACAAATGACACACAGAGCTGCGAGTTTGGGACCTTGGAAACCCACACCGAATATGAGTTC 1500
QY 501 TyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis 520
DB 1501 TACGTGGTGGCTACTCCAGCTGGGGGCCAGCCGAACTCCAGCCAGCCCTGGTGCAAT 1560
QY 521 ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer 540
DB 1561 ACACGTGGACGATGTCCCGAGCGCAGACCCAGCTTACTTGTTCAGGCCCAACCCCTCG 1620
QY 541 AspIleArgValAlaTyrProLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLys 560
DB 1621 GACATCAGGTGGCATGTGCTGCCCTTCCAGCTGAGCAATGGACAGTGTCTGAG 1680
QY 561 TyrIleIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly 580
DB 1681 TACAAGATAGAGTAGCGTTTGGGGAAGGAGATCAGGTTTCTCCACCGAGTGCCTGGA 1740
QY 581 AsnGluThrGlnLeuThrIleuAsnSerLeuGlnProAsnLysValTyrArgValArgIle 600
DB 1741 AATGAGACACAACTTACGTATAACTCACTCAGGCCAAACAAAGTGTACCGAGTCCGGATT 1800
QY 601 SerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTyrMetGlnHisArgThrPro 620
DB 1801 TCAGCTGGCACTGGCGCTGTGCTATGGAGTGCCTTCTCAGTGGATGACAGACAGACCT 1860
QY 621 GlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys 640
DB 1861 GGTGTGCACACAGAGCCAGTGTCTTCCCTTGGCCCTGCAGAAATGGAAGTGGGGCAAG 1920
QY 641 MetGluSerLeuValValSerTyrGlnProProHisProThrGlnIleSerGlyTyr 660
DB 1921 ATGGAGTCCCTGGTGTGTATGGCAGCGCGCCCTCACCCACCCAGATCTCTGGATAC 1980
QY 661 LysLeuTyrTyrGlyValGlyThrGluGluAlaAspGlyAspArgProProGly 680
DB 1981 AAACCTTACTGGGAGAGGTGGAAACAGAGGAGGAGGAGTGGTGGACCGCCCCCAGGG 2040
QY 681 GlyArgGlyAspGlnAlaTyrAspValGlyProValArgLeuLysLysValLysGln 700
DB 2041 GGTGTGGAGATCAGCTTGGGACGTGGGCCCTGGCGCTGGAAGAAAGTGAAGCAG 2100
QY 701 TyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsn 720
DB 2101 TATGAACCTGACCCAGTTCCTTGGCAGGCGGTACGAGGTGAGCTCGTAGCTTTCAAC 2160
QY 721 LysHisGluAspGlyTyrAlaAlaValTyrLysGlyLysThrGluLysAlaProThrPro 740
DB 2161 AAAACAGAGGACGGCTACGCTGCTGTGGAAAGGCAAGAGGAGGAGCGCCACGCGCA 2220
QY 741 AspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsn 760
DB 2221 GACCTGCCATATCCAGAGGGGGCCACCGCTGCTCTCCCGCATGTCCAGCAGAGTCAAC 2280
QY 761 SerSerThrSerIleTyrLeuArgTyrLysLysProAspPheThrThrValLysIleVal 780
DB 2281 AGCTCCACTTCATTTGCTTCCGTGGGAAGACGACACTTACCACTGTCAAGATTGTC 2340
QY 781 AsnTyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThrTyrTyr 800
DB 2341 AACTACACTGTACGCTTCCGCCCTCGGGGCTCAGAAATGCTTCCCTGGTCACTACTAT 2400
QY 801 ThrSerSerGlyGluAspIleIleuIleGlyLeuLysProPheThrLysTyrGluPhe 820
DB 2401 ACCAGCTCTGGAGAAGACATTTCTCATTTGGCGGCCCTGAAACCACTTTACCAAGTACGAGTTT 2460
QY 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840

DB 2461 GCGGTACAGTCCCAAGGAGTGGATATGGATGGGGCCCTTTGGCTCCGTAGAACGCTCC 2520
QY 841 ThrLeuProAspArgProSerThrProSerAspLeuArgLeuSerProLeuThrPro 860
DB 2521 ACCCTGCTGACCGGCTTCAACACCTCTTCTGACCTGGCCCTGAGGCCCTGACACCA 2580
QY 861 SerThrValArgLeuHisTyrProProProThrGluProAsnGlyGluIleValGluTyr 880
DB 2581 TCCACCGTTCCGTTTACATGTGTCTCCCAAGGAGCCCAATGGTGGAGATTGTGGAGTAT 2640
QY 881 LeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeuThrThr 900
DB 2641 CTAAATCTCTACAGCAACACACACCCAGCCCAACACCAAGTGGACACTCTCTCACCACA 2700
QY 901 GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
DB 2701 GAGGGAAACATCTTCAGTGCAGAGTCCATGGCCCTAGAGAGTGACACTCGGTATTTCTTC 2760
QY 921 LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle 940
DB 2761 AAGATGGAGCGCACAGAGTGGGCTGGGCTTCTCCGCTTTCAGGATGTGATT 2820
QY 941 ThrLeuGlnGluThrPheSerAspSerIleuAspValHisAlaValThrGlyIleIleVal 960
DB 2821 ACTCTGCAAGAGACATTTCTCAGACTCTCTTGGATGTGACGCGCTCAAGGGCATCATCGT 2880
QY 961 GlyValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
DB 2881 GGTGTCTGCCCTGGGCTTCTCTGCTCTTGGCTGTGCATGTGTGCTTACGACAAAGC 2940
QY 981 SerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeu 1000
DB 2941 TCCACAGGGAAGCCCTTCCCGGATTGTCTCTCAGGCACCCCGAGGAACCCAGGCTC 3000
QY 1001 TyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeu 1020
DB 3001 TACACAAGAGCTCGGCTTGGGCTCCAGAGTGTCTGCTGCTGCCATGAGTTGGAGTCCCTC 3060
QY 1021 ValHisProArgProGlnAspTyrProProSerProSerAspValGluAspLysAlaGlu 1040
DB 3061 GTGCATCTCTGCTCCAGGATTTGGTCCCAACCTCAGATGTGAAGAAGAGGCTGAA 3120
QY 1041 ValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIle 1060
DB 3121 GTACACAGCCTTATGGGTGGCAGTGTTCAGATTGCCGGGCCACTCCACAGAGAAGATC 3180
QY 1061 SerTrpAlaGlnAlaGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuPro 1080
DB 3181 TCCTGGGCTCAGGCGGGGACCAAACTGGGCAAGCTCTGGGAGGCTCTGGGAGCTGCGCC 3240
QY 1081 GlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
DB 3241 CAGGGTATGTGTCCAAAGCGCGCTCTGACCGCTCTCTGCTGCCCTCCAGCGGAAACCGGG 3300
QY 1101 GlnThrLeuLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLysLys 1120
DB 3301 CAGACACTGTCTGTGCAAGCCCTGTGTATGACGGCATTAAGAGCAACCGGAGAAAGAG 3360
QY 1121 ProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGly 1140
DB 3361 CCGTCCCGAGCTCAGGAATCAGGTGGAACTGAGGTCAATTGTCCACTCCGACTTCGCT 3420
QY 1141 AlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThr 1160
DB 3421 GCATCCAAAGGATGTCTCTGACTCCACCTCCAGACCTGGAGCCAGAGAAACCACTGACT 3480
QY 1161 AlaGluThrLeuProSerThrSerGlyValAlaValAspLeuSerGlnGlyAlaAspTrpLeu 1180
DB 3481 GCAGAGACTCTGCCCTCCAGCTCTGGAGCTGTGTATCTCTCAAGGAGCAGATGCGCTG 3540
QY 1181 GlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200

```

Db      3541  GGcAGGAGCTGGAGGGTGCcAACCAcCAGTGGCCAGAGAGGCTcACCTGCTTG 3600
Qy      1201  ProGluAlaAseAlaSerCySerCySerAspLeuGlnProSerThrAlaIleGlu 1220
Db      3601  CAGAAGCAGCAGTGCCTCTGCTCTAGAGCTCAGCCAGCAGCTGCTATAGAG 3660
Qy      1221  GluAlaProGlyLysSerCyGlnProLysAlaLeuCyAProLeuThrValSerProSer 1240
Db      3661  GAGGCCCTGGGAAAGCTGCCAGCCAAAGCCCTGTGCTCTACAGTCAGCCCAAGC 3720
Qy      1241  LeuProArgAlaProValSerSerAlaGlnValPro 1252
Db      3721  CTTCCAGGAGGCCCTCTCTCTCTGCTCAGGTCCCC 3756

AF176694 6176 bp mRNA linear ROD 27-APR-2000
Mus musculus neighbor of Punc eil protein (Nope) mRNA, complete
cds.
ACCESSION AF176694
VERSION AF176694.1 GI:7650185
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 6176)
AUTHORS Salbaum, J.M. and Kappen, C.
TITLE Cloning and expression of nope, a new mouse gene of the
immunoglobulin superfamily related to guidance receptors
JOURNAL Genomics 64 (1), 15-23 (2000)
MEDLINE 20175427
PUBMED 10708514
REFERENCE 2 (bases 1 to 6176)
AUTHORS Salbaum, J.M.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1999) The Neurosciences Institute, 10640 John Jay
Hopkins Drive, San Diego, CA 92121, USA
FEATURES
    Location/Qualifiers
        1..6176
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="FVB"
            /db_xref="taxon:10090"
            /chromosome="9"
            /map="band D-E1; 24.59 cR from D9Mit143"
        1..6176
            /gene="Nope"
        1..3759
            /gene="Nope"
            /notes="immunoglobulin superfamily; related to guidance
            receptors"
            /codon_start=1
            /product="neighbor of Punc eil protein"
            /protein_id="AAF65930.1"
            /db_xref="GI:7650186"
            /translators="MARADTGRGLVLTCLLSARGELPLPQSTTVKLSCDREGLQVLI
            LGPQAVLDCTLGATAGPPRTVTSKDDGTVLEHNLHLLPNSGLMSSLPQEDS
            DBEALRWKTVSGSLASHPLGVASQVAVVKKLTLDFSLHSPESQVLENGAR
            FECHTGLPAPITTWKQDQVTPPEPRLITLPMKLIQILDQSDSDAGSYRCVATNSAR
            QRFSSALTVALLRGSLEATRGQDVIIVAAPENTTVSGQNVMECVASADPTPFVSW
            VRDQKPISTDVITVLGRTNLLIASAPRHSGVYVCRANKPLTRDFATAAALRLVLAAP
            AISOAPALSRSTASTARFVCRASGEPRLALHWDGIPLRPNGRVKVOGGGSLVIT
            QICLDAGYQCYVAENSAGTACAAAPLAVVREGLPSAPTRVATATPLSSSVLAWER
            PELHSSQIIGFSLHYQKAGVDNVEYQFVNNDDTLQVRDLPEPNTDIFVVAYSQL
            GASRTSPALVLDVPSAQPQLTSSNPSPDIRVAMLPPLSSLSNGQVLYKIKBYG
            LGKEDQVFSTEVENETQTLNSLPNKVYRIRISAGTGAGYGVPSQWQHRTPGVHN
            QSHVPAPALRYAKMBSLVWSQPPHPTQISGYKLYMGVEGTBEERADGDRPPGGR
            GDOAVDPVPLKXKVKQVELTOLVPCRPEVKLVAPKNHEDGYAAVMKGTKEKPTP
            DLFIQGPPLPPAHVAENSSTSIWLKKKDPFTTKVIVNTVTRGPGWGLNASLVT
            YVTSGEDILIGLKPTPKYEFVAVQSHGVDMPGFGSVVERSTLPDRPSTPSPDLRLS
            PLTPSTVRLHWCFPTPEFTIYILYLSNNHTQPBQHWTLTTEGNIIFSAEVHGLS
            DTRYFFKMGARTBEVGPFSRLQDVITLQETFSDSLVDVHATVGIIVGVCLGLICLLAC

```

misc_feature

1..23

/gene="Nope"

/note="unable to sequence region in mRNA"

BASE COUNT 1316 a 1756 c 1634 g 1470 t

ORIGIN

Alignment Scores:

Pred. No.:	Score:	3.04e-260	Length:	6176
Percent Similarity:	6604.00	Matches:	1252	
Best Local Similarity:	100.00%	Conservative:	0	
Query Match:	100.00%	Mismatches:	0	
DB:	10	Indels:	0	
		Gaps:	0	

US-09-754-997A-2 (1-1252) x AF176694 (1-6176)

Qy 1 MetAlaArgAlaAspThrGlyArgGlyLeuLeuValLeuThrPheCysLeuLeuSerAla 20
 Db 1 ATGGCGCGGGCGGACACGGGCGCGGGCTCTGTGCTGACCTTCTGCTGTGTCTCGCG 60
 Qy 21 ArgGlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCyAAspGluGlyPro 40
 Db 61 CGCGGGAGCTGCCATTGCCCCCAGGAGACAACTGCAAGCTGAGCTGTGATGAGGAGCCC 120
 Qy 41 LeuGlnValIleLeuGlyProGlnGlnAlaValValLeuAspCysThrLeuGlyAlaThr 60
 Db 121 CTGCAAGTGATCTCGGGCTTGAGCAGGCTGTGGTGTGAGCTGCACCTTGGGGGCTACA 180
 Qy 61 AlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGluHis 80
 Db 181 GCTGTGGGCTCCCGACAGGGTGACATGGAGCAGAGATGGAGACACTGTACTAGAGCAT 240
 Qy 81 GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu 100
 Db 241 GAGAACCTGCACCTGCTACCCCAATGGCTCCCTGTGGTGTGCTCACCCTAGAGCAAGAA 300
 Qy 101 AspSerAspAspGluAlaLeuArgIleTrpLysValThrGluGlySerTrpSerCys 120
 Db 301 GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAGGTCACCTAGAGGCGAGCTATTCTCT 360
 Qy 121 LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThr 140
 Db 361 CTGGCCCAACAGCCCGCTAGAGTGTGGCCAGCAGGTGTGCTGTGCTCAAGCTTGCCACA 420
 Qy 141 LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg 160
 Db 421 CTCGAAGACTTCTCTCTGCAACCCCGAGTCCCAGATTGTGGAGGAGAAACGGACAGCACGC 480
 Qy 161 PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnVal 180
 Db 481 TTTGAATGTCACACCAAGCGCTTCAGCCCCCATCATTTACTTGGGAAAGGACCAAGGTG 540
 Qy 181 ThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuGlnIleLeuAsp 200
 Db 541 ACCGTGCTTGAGAGAGCCCCGGCTCATCACTCTCCCAAGTGGCTCTCTCAGATCTCTAGAT 600
 Qy 201 ValGlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArg 220
 Db 601 GTCCAGGACAGTGTGAGGCTCTTACCGTGGTGGCCACCAATTCAGCCCCCAACGA 660
 Qy 221 PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly 240
 Db 661 TTCAGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACAGGGGG 720
 Qy 241 GlnAspValValIleValAlaAlaProGluAenThrThrValValSerGlyGlnAenVal 260
 Db 721 CAGGATGTGGTCAATTTGTGGCAGCCCCCAGAGAACACCGGTAGTGTCTGACAGAGATGTA 780

QY	261	ValMetGluCysValAlaSerAlaAspProThrProPheValSerTyrValArgGlnAsp	280	Db	1861	GGTGTGCACACAGCCATGTTCCCTTTGGCCCTGCAGAAATGAAGGTGAGGCGCAAG	1920
Db	781	GTGATGGAGTGGTGGCTGTCTGACCCCACTTTGTGTCTCTGGGTCCGACAGAT	840	QY	641	MetGluSerLeuValSerTyrGlnProProHisProThrGlnIleSerGlyTyr	660
QY	281	GlyLysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSer	300	Db	1921	ATGGAGTCCCTGGTGGTGTATGGCAGCGCCCTCACCCACCCAGATCTCTGGATAC	1980
Db	841	GGAAAGCCTATCTCCACCGATGTATGTTCTGGGCGGAGCCATCTACTATGCCGACG	900	QY	661	LysLeuTyrTyrGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProGly	680
QY	301	AlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAsp	320	Db	1981	AAACTCTACTGGGGAGAGGTGGAAACAGAGGAGGAGGAGATGTGTGACCGCCCCCAGG	2040
Db	901	GGCAGCCTCGGCATCTGGAGTCTATGTCTGCCGAGCCAAAGCCCTCAGCGGTGAC	960	QY	681	GlyArgGlyAspGlnAlaTyrAspValGlyProValArgLeuLysLysValLysGln	700
QY	321	PheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro	340	Db	2041	GGTGTGGAGATCAAGCTTGGACGTGGGGCCCTGGCGCTGAAGAAGAAAGTAGTGAG	2100
Db	961	TTCCGCACCTGGCGCTGTGAGCTCCGAGTGTCTGTGCCCCAGCCATCTCGCAGCACCC	1020	QY	701	TyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsn	720
QY	341	GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu	360	Db	2101	TATGAACTGACCCAGTTAGTCCCTGGCAGCGCTACGAGGTGAAGCTCGTAGCTTTCAAC	2160
Db	1021	GAGGCGCTCTCCGGAAGGGGCGCAGACCGCGCGCTTCGTGTCCGGCGCTCCGGGAG	1080	QY	721	LysHisGluAspGlyTyrAlaAlaValTyrLysGlyLysThrGluLysAlaProThrPro	740
QY	361	ProArgProAlaLeuHisTyrLeuHisAspGlyIleProLeuArgProAsnGlyArgVal	380	Db	2161	AAACACGAGGACGGCTACGCTGTGTGTGAAGGCAAGACGAGAGAGGCGCCACGCCA	2220
Db	1081	CCACGCGCGCGCTGCACCTGGCTGCACGCGGATCCCGTTGGACCCCAATGGGCGCTC	1140	QY	741	AspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsn	760
QY	381	LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly	400	Db	2221	GACCTGGCTATCCACAGGGGGCCACCGCTGCCTCTGCCATGTGTCCACGCAAGTCAAC	2280
Db	1141	AAGGTGCAGGCGGTGGCGGACGTGTGTCTCATCTCAGATCGGCTCGCAGGACCTGGC	1200	QY	761	SerSerThrSerIleTyrLeuArgTyrLysLysProAspPheThrValLysIleVal	780
QY	401	TyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAla	420	Db	2281	AGCTCCACTTCCATTGGCTTCGGTGGAGAGCCAGACTTTACACTGTGONAGATTGTC	2340
Db	1201	TACTACCACTGGTACGAAACACAGCGGGAACTGGCTGTGGCGCTGGCGCCCTGGCG	1260	QY	781	AsnTyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThrTyrTyr	800
QY	421	ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer	440	Db	2341	AACTACACTGTATGCTTGGCGCCCTGGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT	2400
Db	1261	GTAGTGGTGGCGAGGGCTGGCCAGCGCCCGACTCGGTCACAGCCACGCGCTGAGC	1320	QY	801	ThrSerSerGlyGluAspIleLeuIleGlyLysLysProPheThrLysTyrGluPhe	820
QY	441	SerSerSerValLeuValAlaTyrGluArgProGluLeuHisSerGluIleIleGly	460	Db	2401	ACCAAGCTCTGGAGAAGACATTTCTCATTTGGCGGCTGAAACCATTTTACCAAGTACAGT	2460
Db	1321	AGCTCCTCTGTGTGTGGTGGCGGCGGCTGAGTTGCAAGCGAGCAATCATTTGGC	1380	QY	821	AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer	840
QY	461	PheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaVal	480	Db	2461	CGGTACAGTCCCAAGGAGTGGATATGATGGGCGCTTTGGCTCCGTGAGAACGCTCC	2520
Db	1381	TTCTCTCTTCACTACCAAAAGCAAGGGAGTGGACAAATGTGGAGTACCAATTTGCA	1440	QY	841	ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro	860
QY	481	AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPhe	500	Db	2521	ACCTGGCTGACCGGCTTCAACACTCTCTCTGACCTGGCGCTGAGCCCTCTGACACCA	2580
Db	1441	AACATGACACACAGAGCTGCGAGTTGGGACCTTGGAAACCCACACCGATTATGATTC	1500	QY	861	SerThrValArgLeuHisTyrCysProProThrGluProAsnGlyGluIleValGluTyr	880
QY	501	TyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis	520	Db	2581	TCCACCGTTCGGTTTACACTGGTGTCCCGCCACGAGGCGCAATGGTGAAGATTGTGAGTAT	2640
Db	1501	TACGTGGTGGCTACTCCAGCTGGGGGCGAGCGCACTCCAGCCAGCCCTGGTGCAT	1560	QY	881	LeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeuThrThr	900
QY	521	ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer	540	Db	2641	CTAATTTCTACAGCAACACACACCGCCGCAACACCAAGTGGACACTCTCTCACCA	2700
Db	1561	ACACTGGACGATGTCCCGCAGCAGCAGCCCACTTACCTTGTTCAGCCCAACCCCTCG	1620	QY	901	GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe	920
QY	541	AspIleArgValAlaTyrLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLys	560	Db	2701	GAGGGAACATCTTCAGTGCAGAGGTCCATGGCTTAGAGAGTGACACTCGGTATTTCTTC	2760
Db	1621	GACATCAGGTGGCATGGCTGCCCCCTGCGCTCCAGCTGAGCAATGGACAGGTGCTGAAG	1680	QY	921	LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle	940
QY	561	TyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly	580	Db	2761	AAGATGGAGCGCCACAGAGGTGGGCGCTGGGCGCTTTTCCGCTTGCAGATGTGATT	2820
Db	1681	TACAGATAGAGTACGGTTTGGGAAGGAAGATCAGGTTTTTCTCCACCGAGTGTCTGA	1740	QY	941	ThrLeuGlnThrThrPheSerAspSerLeuAspValHisAlaValThrGlyIleVal	960
QY	581	AsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIle	600	Db	2821	ACTCTGCAAGAGACATTTCTCAGACTCTCTGATGTGACCGCGCTCAACGCGGATCATCTCGTG	2880
Db	1741	AATGAGACACAACTTACGTAACTCACTTACGCCAAACAAAGTGTACCGAGTCCGAT	1800	QY	961	GlyValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGlnSer	980
QY	601	SerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTyrMetGlnHisArgThrPro	620	Db	2881	GGTGTCTGGCGCTCTCTGCTCTGCGCTCTGATGTGTGTGTGTGTGTGTGTGTGTGTGT	2940
Db	1801	TCAGCTGGCAGCTGGCGCTGTGATGAGTCCCTTCTCAGTGGATGATGACGACGACCT	1860	QY	981	SerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeu	1000
QY	621	GlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys	640				

```

Db      2941  TCCACAGGGAAGCCCTTCCCGATTGTCCTCTCAGGCACCCAGGAACCCAGCGCTC 3000
Qy      1001  TyThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeu 1020
Db      3001  TACACAGAGCTCGGCTTGGCCCTCCAGTGTCCCTGCTGCCCATGATGGAGTCCCTC 3060
Qy      1021  ValHisProArgProGlnAepTrpSerProProSerAspValGluAepLysAlaGlu 1040
Db      3061  GTGCATCTCGTCCCGAGATTGGTCCCAACACCCCTCAGATGTGGAAGACAGGCTGAA 3120
Qy      1041  ValHisSerLeuMetGlySerValSerAspCysArgGlyHisSerLysArgLysLeu 1060
Db      3121  GTACACAGCCCTTATGGGTGGCAGTGTTCAGATTGCCGGGGCCACCTCCCAAGAGAAGATC 3180
Qy      1061  SerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuPro 1080
Db      3181  TCTGTGGCTCAGGACGGGACCAACTGGGCAAGCTCTCTGGGCAAGCTGTGAGCTGCC 3240
Qy      1081  GlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
Db      3241  CAGGAGTAGTGGTCCAAAGCCGGCTCTGACCCGCTGCTCTGCTGCCCTCAGCGGAACCGG 3300
Qy      1101  GlnThrLeuLeuGlnAlaLeuValTyrAspGlyLysSerLysArgLysLys 1120
Db      3301  CAGACACTGCTGCTGCAAGCCCTGGTGTATGACGGCATTAAGAGCAACCGGAGAAAGAG 3360
Qy      1121  ProSerProAlaCysArgGlnValGluAlaGluValLeuValHisSerAspPheGly 1140
Db      3361  CGTCCCGACCTGCAGGAATCAGGTGGAGAGCTGAGTCTATGTCACCTCCGACTTCG 3420
Qy      1141  AlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThr 1160
Db      3421  GCATCCAAAGGATGTCTGACCTCCACCTCCAAAGACCTGGAGCCAGAGGAACCACTG 3480
Qy      1161  AlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeu 1180
Db      3481  GCAGAGACTGCTGCTTCCAGCTGGAGCTGTGATGCTGTCTCAAGAGCAGACTGGCTG 3540
Qy      1181  GlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
Db      3541  GGAGGGAGCTGGAGAGGTGCCAACCAACAGCTGGCCAGAGGCTCACCCTGCTTG 3600
Qy      1201  ProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAlaLeuGlu 1220
Db      3601  CCAGAGCAGCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3660
Qy      1221  GluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSer 1240
Db      3661  GAGGCCCTCGGAAAAGCTGCCAGCCCAAGCCCTGTGTCTCTTAACAGTCAGCCCAAGC 3720
Qy      1241  LeuProArgAlaProValSerAlaGlnValPro 1252
Db      3721  CTTCCAGGGCCCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3756

RESULT 4
LOCUS   AB052620
DEFINITION Mus musculus DDM36 mRNA, complete cds.
ACCESSION AB052620
VERSION   AB052620.1 GI:11862938
KEYWORDS
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Murakami,H., Nakayama,T., Yamamoto,H., Hosaka,T., Aoyama,T., Nagayama,S., Oka,M., Kiyono,T., Sasaki,M.S., Nakamura,T. and Toguchida,J
TITLE   Up-regulation of a ras effector and down-regulation of a cell adhesion molecule are associated with transformation of osteoblasts
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 6219)

```

AUTHORS TITLE JOURNAL

Toguchida,J., Nakamata,T., Murakami,H., Nakayama,T. and Nakamura,T.
Direct Submission
Submitted (11-DEC-2000) Junya Toguchida, Kyoto University,
Institute for Frontier Medical Sciences; 53 Kawahara-cho, Shogoin,
Sakyo-ku, Kyoto, Kyoto 606-8507, Japan
(E-mail:tcogjun@frontier.kyoto-u.ac.jp, Tel:81-75-751-4134,
Fax:81-75-751-4144)

FEATURES source

Location/Qualifiers
1..6219
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6BA"
/db_xref="taxon:10090"
/clone="DDM36"
/cell_line="MMC2"
/cell_type="osteoblast"
/note="similar to Nope (NM_020043)-vector:pDDM36ORP"

gene

1..6219
/gene="DDM36"

CDS

50..3808
/gene="DDM36"
/note="similar to Nope (NP_064427)"
/codon_start=1
/evidence=experimental
/product="DDM36"
/protein_id="BAB19278.1"
/db_xref="GI:11862938"
/translation="MARADTGRLLVLTFCLLSARGELPLQETTVKLSCDGGLQVILGPEQAVLDCTLCATAGPTVTWVSKDGTVLLEHNLHLLPNSLWLSPLQEDSDEERLRWKVTEGYSCLAHSPLGVAQVAVKLTATLDFSLPESQIVBENGSTARFECHTGLUPAPIITWKKQVTPRESRLITLPNGVLQILDVQSDAGSRVATNSARVRQGEASLTVALRGSLTADQDVIVAPENTTVSGSVQVWECVASADBLRVLPVSWVRQDKPTSTDIVILGRTNLIAAQPRHSGVYVCRANKPRTRDPATAAELRLVLAAPAISSOAPALSRTRASTFVCRASGEPALHLDGIPLRNGRVKVGQGGGLVITQIGLDQGYOCVAENSAGTACAAPLAVVREGLPSAPTRVATPLSSSVLVAWERFELHSEQIIGSLHYQKARGVDNVEYQFANNDTTELQRLDLENTDTEFYVVAYSQLGASRTSPALVHTLDDVPSAQLTNSPNPSDINWMLPLPSSLSNGQVLKYLIEYLGKEDQVSTFVPGNETQTLNLPKNKRVIRISAGTGGVGPVQWQHRTPGVHNSHNSHTPEHOMTLTTLTGNIPSAVHGLSDTRYPFGKARTFVGPPSRQDVITLQKTFSDSLDHVATGIVLGVCLCLLACMCAGLRSHREALPGLSSGTPGNALYTRALGPPSPVPAHLESLVHPRPQDMSPPSDVEDAEVHSLMGSSVDCRSHKRSIWAQAGPNWAGSWAGCPLPGSGSPRALTRALLPAGTGTLLQALVYDAIKNGRKEPSACRQVBEAVIYHSDFGASKGCPDUHLQDLRPPSPPLTARTLPSTGSAVDLSQADWLGRGLGCGQPTTSGPERLTCLEPAA

BASE COUNT 1319 a 1772 c 1652 g 1476 t

Alignment Scores:

Pred. No.: 2,63e-257 Length: 6219
Score: 6532.00 Matches: 1241
Percent Similarity: 99.44% Conservative: 4
Best Local Similarity: 99.12% Mismatches: 7
Query Match: 98.91% Indels: 0
DB: 10 Gaps: 0

US-09-754-997A-2 (1-1252) x AB052620 (1-6219)

Qy 1 MetAlaArgAlaAspThrGlyArgGlyLeuLeuValLeuThrPheCysLeuLeuSerAla 20
Db 50 ATGGCCGGCGGCGGACACCGGGCCCGGGCTCTCTGTGTGTGACTCTTCTGCTGTCTGCGG 109
Qy 21 ArgGlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCysAspGluGlyPro 40
Db 110 CGCGGGGAGCTGCCATTGCCCCCAGGAGACAACTGTCAAGCTGAGCTGTGTATGAGGACCC 169
Qy 41 LeuGlnValLysLeuGlyProGluGlnAlaValLeuAspCysThrLeuGlyAlaThr 60
Db 170 CTGCAAGTGTATCTGGGCCCTGAGCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 229

QY 61 AlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGluHis 80
DB 230 GCTGCTGGGCTCCACACAGGGTGACATGGAGAGGATGGAGACACTGTACTAGAGCAT 289
QY 81 GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu 100
DB 290 GAGAACCTGGCACTGCTACCCAAATGGCTCCCTGTGGCTGTCTCTACCCCTAGAGCAAGAA 349
QY 101 AspSerAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTrpSerCys 120
DB 350 GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAAGTCACTGAGGGCAGCTATTCTCTGT 409
QY 121 LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThr 140
DB 410 CTGGGCCACAGCCCTAGGAGTGGTGGCCAGCCAGGTGTGCTGTGGTCAAGCTTGCCACA 469
QY 141 LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg 160
DB 470 CTCGAAGACTTCTCTGCAACCCGAGTCCCAAGATTGGAGAGAACGGACAGCAGC 529
QY 161 PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnVal 180
DB 530 TTGGAATGCCACACCAAGGCTCTCCAGCCCCCATATTACTTGGAAAGACCAAGGTG 589
QY 181 ThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAsp 200
DB 590 ACCGTGGCTGAGAGTCCGGCTCATCTCTTCCCAATGGCGTCTCCAGATCCTAGAT 649
QY 201 ValGlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArg 220
DB 650 GTCCAGGACAGTGAAGAGGCTCTTACCCTGCGTGGCCACCAATTCAGCCCGCCCAACGA 709
QY 221 PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly 240
DB 710 TTCACCCAGAGGCTCTGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAAGGGG 769
QY 241 GlnAspValValIleValAlaAlaProGluAsnThrThrValSerGlyClnAsnVal 260
DB 770 CAGGATGTGGTCAATTGTGGCAGCCCGCAGAGAACACACCGTAGTGTCTGGACAGAGTGA 829
QY 261 ValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp 280
DB 830 GTGATGGAGTGGTGGCTCTGCTGACCCCACTTGTGTCTGGTCCGACAGGAT 889
QY 281 GlyLysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSer 300
DB 890 GGAAAGCCTAATCTCCACGGATGTCACTGTTCTGGGCCCGGACCAATCTACTCATGCCAGC 949
QY 301 AlaGlnProArgHisSerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAsp 320
DB 950 GCGCAGCCTCGGCACCTCTGGAGTCTATGTCTGCCGAGCCAAAGCCCGCAGCGGTGAT 1009
QY 321 PheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro 340
DB 1010 TTCGCACTCGCGGCTGTGAGCTCCAGTGTCTGTGCCCCAGCCCATCTCGCAGCGGCC 1069
QY 341 GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyCln 360
DB 1070 GAGGCGCTCTCGGACCGGGCCAGCACCGCGCGCTTCTGTGTCCGGGCGTCCGGGAG 1129
QY 361 ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgVal 380
DB 1130 CCACGGCCCGCTGCACTGGCTGACACCGGGATCCGTTGGACCCCAATGGGCGCTC 1189
QY 381 LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly 400
DB 1190 AAGGTGACGGGCGTGGGGGAGCTTGGTTCATCTACATCGGCGCTTTCAGGACGCTGGC 1249
QY 401 TyrTrpGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAla 420
DB 1250 TACTACAGTGTGATAGCAAAACAGCGGGAACTGCTCTGTGGCGTGGCGCCCTGGCG 1309

QY 421 ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer 440
DB 1310 GTAGTGTGGCTGGAGGGGCTGCCAGCGCCCCGACTCGGGTCAACAGCCACCCCTGAGC 1369
QY 441 SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGly 460
DB 1370 AGCTCTCTGTGCTGTGGCTGGAGCGGCTGAGTTGGACACGGAGCAATCATTTGGC 1429
QY 461 PheSerLeuHisTrpGlnLysAlaArgGlyValAspAsnValGluTrpGlnPheAlaVal 480
DB 1430 TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAAATGTGGAGTACCACTTTGCAGTA 1489
QY 481 AsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPhe 500
DB 1490 AACAAATGACACACAGAGCTGAGTTGGGACCTTGGAAACCAACACGGATTTATGAGTTC 1549
QY 501 TyrValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis 520
DB 1550 TACGTGTGGCTACTCTCCAGCTGGGGCCAGCCGAACCTCCAGCCAGCCCTGGTGAT 1609
QY 521 ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer 540
DB 1610 ACCTGGAGCATGTCCCAGCGCAGCACCCAGCTTACCTTGTCCAGCCCCCAACCCCTCG 1669
QY 541 AspIleArgValAlaTrpLeuProLeuProSerSerSerLeuSerAsnGlyGlnValLeuLys 560
DB 1670 GACATCAGGGTGGATGGCTGCCCTCTCCAGCTGGAGCAATGGACAGCTGCTGAAG 1729
QY 561 TyrLysIleGluTrpGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly 580
DB 1730 TACAAATAGAGTACGGTTTGGGAGAGAGATCAGGTTTTCTCCAGGAGTGGCTGGA 1789
QY 581 AsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTrpArgValArgIle 600
DB 1790 AATGAGACACACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGAGTT 1849
QY 601 SerAlaGlyThrGlyAlaGlyTrpGlyValProSerGlnTrpMetGlnHisArgThrPro 620
DB 1850 TCAGCTGGCACTGGCGCTGGCTATGAGTGCCTTCTCAGTGGATGTCAGCACAGCACCT 1909
QY 621 GlyValHisGlnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys 640
DB 1910 GGTGTGCACAAACAGAGCCATGTTCCCTTGTGCCCTCAGAAATGAAAGGTGAGGGCAAG 1969
QY 641 MetGluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTrp 660
DB 1970 ATGGAGTCCCTGGTGGTGTCAATGGACGCGCCCTCACCCCAACAGATCTCTGGATAC 2029
QY 661 LysLeuTrpTrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProProGly 680
DB 2030 AAACCTCTACGGAGAGGTGGGAAACAGAGAGGAGGAGATGGTGCACGCCCCCAGGG 2089
QY 681 GlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGln 700
DB 2090 GGTGTGGAGATCAAGCTTGGGACGTGGGCGCTGGCGGTGGAAGAAAGTGAAGCAG 2149
QY 701 TyrGluLeuThrGlnLeuValProGlyArgProTrpGluValLysLeuValAlaPheAsn 720
DB 2150 TATGAACTGACCCAGTTAGTCCCTTGGCAGGCTGTACAGGTTGAAGCTCTGTAGCTTTCAAC 2209
QY 721 LysHisGluAspGlyTrpAlaAlaValTrpLysGlyLysThrGluLysAlaProThrPro 740
DB 2210 AAACACAGAGACGGCTACCTGCTGTGTGGAGGCAAGACGAGAGAGGCGCCACCGCCA 2269
QY 741 AspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsn 760
DB 2270 GACCTGCTTATCAGAGGGGCGCACCGCTGCTCTGCCCATGTCTCCAGCAGAGTCAAC 2329
QY 761 SerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleVal 780
DB 2330 AGCTCCACTTCCATTTGGCTTCGGTGGAGAGGACAGACTTTACCACTGTCAAGATTGTC 2389
QY 781 AsnTrpThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTrpTyr 800

```

2390 AACTACACTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCCTCGTCACTACTAT 2449
QY ThrSerSerGlyGluAaplleLeuIledlyGlyLeuIysProPheThrIlysTyGluPhe 820
Db ACCAGCTCTGGAGAGACATTCATTCATGGCGGCTGAAACCAITTTACCAAGTAGCAGATT 2509
QY AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840
Db GCGGTACAGTCCCAAGAGTGGATATGATGGGCCCTTTGGCTCCGTCGTAGAACGCTCC 2569
QY ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
Db ACCGTGCTGACCGGCTTCAACACCTCTCTGACCTGGCGCTCAGAGCCCTGACACCA 2629
QY SerThrValArgLeuHisTyrCysProProThrGluProAnGlyGluIleValGluTyr 880
Db TCCACCGCTTCGGTTACACTGGTGTGCCCCCAGGAGCCCAATGGTGAGATTGTGGAGTAT 2689
QY LeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeuThrThr 900
Db CTAATTCCTACAGCAACACCAACCCAGCCGCAACACCACTGGACACTGCTCACCACA 2749
QY GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
Db GAGGAAACATCTTCAGTGCAGAGTCCATGGCTCAGAGTGCAGCTCGGTATTTCTTC 2809
QY LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle 940
Db AAGATGGAGCCGCAAGAGGTGGGCTTGGGCCCTTTTCCCGCTTGCAGAGTGTGATT 2869
QY ThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIleLeVal 960
Db ACTCTGCAAGACATCTTCAGACTCTCTGGATGTGCAGCGCTCAGCGGCATCATGTG 2929
QY GlyValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
Db GGTGTGCTGGGCCCTCTCTGCTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2989
QY SerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeu 1000
Db TCCACAGGGAAGCCCTGCGGATTTGCTCTCAGGCAACCCAGGAAACCCAGCGCTC 3049
QY TyrThrArgAlaArgLeuGlyProProSerValProAlaHisGluLeuGluSerLeu 1020
Db TACACAGAGCTCAGCTTGGGCCCCCAGTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 3109
QY ValHisProArgProGlnAspTyrProProProSerAspValGluAspLysAlaGlu 1040
Db GTGCATCTCTGCTCCCAAGGATTTGCTCCCAACCCCTCAGATGTGGAAGACAAGCTGAA 3169
QY ValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIle 1060
Db GTACACAGCTTATGGGTGGCAGTGTTCAGATTGCGCGGCGCACTCCACAGAGAAGATC 3229
QY SerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuPro 1080
Db TCCGTGGGCTCAGGCGGGGACCAACTGGGCGAGCTCTCGGCGAGCTGTGAGCTGCC 3289
QY GlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
Db CAGGAGTAGTGTCCAGGCGGCTCTGACCGTCTCTGCTGCTCTCAGCGGGAACCGGG 3349
QY GlnThrLeuLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLysLys 1120
Db CAGACACTGCTGCTGCAAGCCCTGGTATATGACGCCATAAAGAGCAACGGGAGAAAG 3409
QY ProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGly 1140
Db CCGTCCCCAGCTCGAGAATCAGGTGAAGCTCAGGTGATTGTGCCACTCCGATTCGGT 3469
QY AlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThr 1160
```

```

3470 GCATCAAAGGATGCTCTGACCTCCACCTCCCAAGACCTGGAGCCAGAGAAACCACTGACT 3529
QY AlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeu 1180
Db GCAGAGACTCTGCTTCCAGCTCTGAGGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTG 3589
QY GlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
Db GGCAGGAGCTGGAGGGTGCCCAACCAACAGCAGTGGGCCAGAGAGGCTCACCTGCTTG 3649
QY ProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAlaIleGlu 1220
Db CCAGAGAGCAGCAGTGCCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3709
QY GluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSer 1240
Db GAGGCCCTGGGAAAGCTGCCAGGCCAAGCCCTGTGTCTCTCTCTCTCTCTCTCTCTCT 3769
QY LeuProArgAlaProValSerSerAlaGlnValPro 1252
Db CTTCCAGGCGCCCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3805

RESULT 5
AB052621
LOCUS Mus musculus DDM368 mRNA, complete cds.
DEFINITION Mus musculus DDM368 mRNA, complete cds.
ACCESSION AB052621
VERSION AB052621.1 GI:11862940
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Murakami,H., Nakamata,T., Nakayama,T., Yamamoto,H., Hosaka,T.,
Aoyama,T., Nagayama,S., Oka,M., Kiyono,T., Sasaki,M.S., Nakamura,T.
and Toguchida,J.
TITLE Up-regulation of a ras effector and down-regulation of a cell
adhesion molecule are associated with transformation of osteoblasts
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 6222)
AUTHORS Toguchida,J., Nakamata,T., Murakami,H., Nakayama,T. and Nakamura,T.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-2000) Junya Toguchida, Kyoto University,
Institute for Frontier Medical Sciences; 53 Kawahara-cho, Shogoin,
Sakyo-ku, Kyoto, Kyoto 606-8507, Japan
(E-mail:togjun@frontier.kyoto-u.ac.jp, Tel:81-75-751-4134,
Fax:81-75-751-4144)
FEATURES
Location/Qualifiers
1..6222
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6BA"
/db_xref="taxon:10090"
/clone="DDM368"
/cell_line="MMC2"
/cell_type="osteoblast"
/notes="similar to Nope (NM_020043)-vector.pDDM368ORP"
1..6222
/gene="DDM368"
50..3811
/gene="DDM368"
/notes="similar to Nope (NP_064427)"
/codon_start=1
/evidence=experimental
/product="DDM368"
/protein_id="BAB19279.1"
/db_xref="GI:11862941"
/translation="MARADTGRGLVLTCLLSARGELPLPQRTTVKLSCDEGPLQV
LGPQAVLDCTGATAAGPPTPTWSKGDVTLEHNLHPNGSLWSSPLEQSDS
DEEALRIWKVTGYSYSLAHPLGVASQVAVKLATLEDLSLHPSESQVATNG
FECHTGLPAPIITWEKQVTPPESLRITLPLNGVLIQDQSDAGSRCTVATNSAR
QRPSQASLTVALRGSLEATRGDQDVIVAAPENTTVVSGSVVMECVASADPTPTVSV

```

VROGKPISTDVIVLGTNLIASQPBHSGVYVCRAKPRTRDPATATAAELRLVLAAP
 AISQPEALSTRASTARFVCRASGEPALHMLHDGILPRLNGKVRKVGSGSLVIT
 QIGLQAGYQCYAENSAGTAAAPLAVVREGLPFSAPTRVATPPLSSSVLVAMER
 PELSHQYIGLSYKARKGVNVEYQFVAVNDTTELOVRLDLEPNTDPTFVAVAYSQL
 GASRTSPALVHTLDDVPSPALQTLSSPNPSDIRVAMLPPLSSLSNGVLYKIEYV
 LKSEDOVSTEVPGNETQTLNLOPNKVRVRIISACTGAGYGVPSOMOHRTPGVH
 NOSHVPAPELKVAPKWSLIVSWQPPHPTQISGYKLYWREVGTEBERADGPRPGG
 RGDQAMVDGVRLLKKVKQYBELTQVLPGLRYEVLVAFNKHEDGTAAVWKGKTKRAPT
 PDLPIQRPPLPPAHVHAESNSTSIWLRKKPDTTVKIVNIVTRPGPWGLRNASLV
 TYTSSGDIILIGLKKPTKYEFVQSHGVMDGPFQSVVERSTLPDRPSTPDDSLR
 SPLTSTVRLHWCPTPENGEIVELYLLYNNHTQPEHOMTLLTTEGNIPSARVHGLE
 SDTRYFFKMGARTVGPFPRLQDVLTLQKTSFDSLNDVHATVGIIVGVCLGLLCLLA
 CMKAGLRSSHREALPGLSSSGTFCNFPALYTRALRGLPPSPAAHELSLHVPRQDWS
 PPSDVEDKAVHSLMGVSDCRGSKKLSMAQAGGPNWAGWAGCELPFGSGGPRP
 ALTRALLPPAGTGTLLQALVDAIKNSGRKPSPCRNQVREAVIIVSHDFPAGSKC
 PDLHLQDLEPEPTTAETLPSTGAVDLSQDAGWMLRELGCQPTTSGRSLTCLPEA
 ASACSCSDLOFSTAIBAPGKSCQPKALCPLTVSPSLPRAPVSSAQP"

BASE COUNT 1321 a 1772 c 1653 g 1476 t
 ORIGIN

Alignment Scores:

Pred. No.: 7,03e-257 Length: 6222
 Score: 6521.50 Matches: 1241
 Percent Similarity: 99.36% Conservative: 4
 Best Local Similarity: 99.04% Mismatches: 7
 Query Match: 98.75% Indels: 1
 DB: 10 Gaps: 1

US-09-754-997A-2 (1-1252) x AB052621 (1-6222)

Qy	1	MetAlaArgAlaAspThrGlyArgGlyLeuLeuValLeuThrPheCysLeuLeuSerAla	20
Db	50	ATGGCGCGGGGACACACGGCGCGGCTCTGGTGTGACCTTCTGCTGCTGCTGCGG	109
Qy	21	ArgGlyGluLeuProLeuProGlnGluThrValLeuLeuSerCysAspGluClyPro	40
Db	110	CGCGGGAGCTGCTGCTATGCCCCGAGAGCACTGTCAAGCTGAGCTGTGATGAGGAGCC	169
Qy	41	LeuGlnValIleLeuGlyProGluGlnAlaValValLeuLeuAspCysThrLeuGlyAlaThr	60
Db	170	CTGCAAGTGATCTGGGCGCTGAGCAGGCTGTGGTGTGACTGTGACTTTGGGGCTACA	229
Qy	61	AlaAlaGlyProProThrArgValThrTrpSerIysAspGlyAspThrValLeuGluHis	80
Db	230	GCTGTGGGCTCCGACAGGGTGACATGAGCAAGGATGGAGACACTGTACTAGAGCAT	289
Qy	81	GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu	100
Db	290	GAGAACCTGACCTGCTACCCAAATGGCTCCCTGTGGCTGTCTCCTACCCCTAGAGCAAGAA	349
Qy	101	AspSerAspGluGluAlaLeuArgIleTrpIysValThrGluGlySerTyrSerCys	120
Db	350	GACAGCGATGATGAGAGAGCTCTTAGGATCTGGAAAGTCACTAGGCGGAGCTATTCTCTGT	409
Qy	121	LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLeuAlaThr	140
Db	410	CTGGCCCAACAGCCCTAGGAGTGTGGCCAGCCAGGTTGCTGTGCTCAAGCTTGCCACA	469
Qy	141	LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg	160
Db	470	CTCGAAGACTTCTCTCTGCACCCCGAGTCCCAGATTGTGGAGGAGAACCGGACACAGCC	529
Qy	161	PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluIysAspGlnVal	180
Db	530	TTTGAAATGCCACACCAAGGCGCTTCCAGCCCCCATCATTTACTTGGGAAAGAGACAGGTGT	589
Qy	181	ThrValProGluGluProAspGluIleThrLeuProIysTrpLeuLeuGlnIleLeuAsp	200
Db	590	ACCGTGCCTGAGGAGTCCGGGCTCATCATCTCTCCCAATGGCGTCTCCAGATCTTAGAT	649
Qy	201	ValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArg	220
Db	650	GTCCAGGACAGTGATGACAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAAGCA	709

Qy	221	PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly	240
Db	710	TTTACGAGGAGGCGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAAGGGG	769
Qy	241	GlnAspValValIleValAlaAaProGluAsnThrThrValValSerGlyGlnAsnVal	260
Db	770	CAGATGTGGTCATTGTGGCAGCCCAAGAGAACCCAGGTAGTGTCTGGACAGAGTGTA	829
Qy	261	ValMetGluCysValAlaAaSerAlaAspProThrProPheValSerTrpValArgGlnAsp	280
Db	830	GTGATGAGTGGTGGGCTCTGCTGAGCCCAACCCCTTTTGTGTCTGGTCCGACAGAT	889
Qy	281	GlyIysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSer	300
Db	890	GGAAAGCTATCTCACCGATGTCTGCTTCTGGCGCGACCAATCTACTCATGCCAGC	949
Qy	301	AlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnIysProLeuThrArgAsp	320
Db	950	GCGCAGGCTCGGCACCTCTGGAGTCTATGTCTGCCGAGCCAAAGCCCGCCAGCGGTAT	1009
Qy	321	PheAlaThrAlaAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro	340
Db	1010	TTCCGCACTGGGCTGTGAGCTCCGAGTGTCTGCTGCCCGCCAGCCATCTCCGAGGCGCC	1069
Qy	341	GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu	360
Db	1070	GAGCGCTCTCGCGAGCGCGGCCAGCACCGCGCTCTGTTGTGCGCGCTCCGCGGAG	1129
Qy	361	ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgVal	380
Db	1130	CCACGGGCGCGCTGTCACCTGGCTGCACGAGGATCCGTTGCGACCCCAATGGGCGCGCTC	1189
Qy	381	LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly	400
Db	1190	AGGTGAGGGCGGTGGCGGAGGTGTGTCTACTCATCTCATGATCGGCTTCAGAGAGCTGCG	1249
Qy	401	TyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAla	420
Db	1250	TACTACAGTGCCTAGCAGAAAAACGCGCGGAACTGCTGTGCTGCGCTCGCCCTCGCG	1309
Qy	421	ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer	440
Db	1310	GTAGTGTGGCGAGGGGCTGCCAGCGCCCGCTCGGGTCAACAGCCACCGCTGAGC	1369
Qy	441	SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGly	460
Db	1370	AGCTCTCTGTGCTGTGGCTGGAGCGGCTGAGTTGCACAGAGCAATCATTTGCG	1429
Qy	461	PheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaVal	480
Db	1430	TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGCAATGTGGAGTACCAAGTTTGCACTA	1489
Qy	481	AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPhe	500
Db	1490	AACATGACACACAGAGCTGAGGTTCCGGACCTCGGAACCAACACACGATATAGTTC	1549
Qy	501	TyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis	520
Db	1550	TACGTGTGGCTTACTCCAGCTGGGGCGGAGCCGCAACCTCCAGCCCGGCTGTGTGAT	1609
Qy	521	ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer	540
Db	1610	ACACTGGACGATGTCTCCCGAGCGGACACCCAGCTTACTTGTTCAGGCCCCCAACCCCTCG	1669
Qy	541	AspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuIys	560
Db	1670	GACATCAGGTTGGATGCTGCTGCCCTCCAGCTTCCAGCAATGGACAGGTGCTGAG	1729
Qy	561	TyrIysIleGluTyrGlyLeuGlyLys---GluAspGlnValPheSerThrGluValPro	579
Db	1730	TACAAGATAGAGTACGGTTTGGGGAGGAAGAGATCAGGTTTTCTCCACCGGAGTGCCT	1789

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Aizawa, H., Yuasa, S., Nakajima, D., Nagase, T., Ohara, O. and Koga, H.
 Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries of DNA Res. 10, 35-48 (2003)
 2 (bases 1 to 6301)
 Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.
 Direct Submission
 Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: mouse@kazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)
 The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing: Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.
 Location/Qualifiers

FEATURES

source

1. .6301
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="mbg00013"
 /issue_type="brain"
 /dev_stage="adult"
 /note="vector:modified pBC SK+"
 1. .6301
 /gene="mKIAA1628"

gene

<1. .3887
 /gene="mKIAA1628"

CDS

/note="CDS is predicted by in silico analysis. Start codon is not identified."
 /codon_start=3
 /evidence=not_experimental
 /product="mKIAA1628 protein"
 /protein_id="BAC65817.1"
 /db_xref="GI:28972802"

/translation="EGSQNPANWGHGRLPLPORTTVKLSCEGQLQVILGPQAVV
 LQCTGATAGAPTRVYNSKGDVTLHENLHLLNGSLSPLEQEDSDDEALRI
 WKTFGSGSLAHPGLVAVASQVAVKATLEDFSLHPSQIVYENGTAFFCHTKGL
 PAPIITWEDQVTPRESRLITLPGVLIQIDQSDAGYRCVATNGARQRFQSRAS
 LTVALRGLEATRGVDQVIVAPENTVVGOSVVMBCVASADPTPFVSWVRQGSQK
 KELGIOTSCDAALQGCNGLVVGEGVGTVPASPSILLCHOMESLSPMSFWAG
 PLYSPARSLGLENSAFTSPARVISPLRLISSECLLPQPSRRPRSRGRGAPR
 ASCAGPSHGPRCTCTTCGDCDPMGRVKVQGGSLVITQIGLDAGYVOCVAENS
 AGTCAAPALAVVREGLSPATRVATPTELSSSVLWAPERPELHSEQIIGFSLHYK
 ARGVONVEYQFVANDTLEQVLDLEPNTDYFYVAYSQLASRTSSPALVHTLDDV
 PSAAPOLTLSSPNPSDIRVAMLPLSSLSNGOVLYKIEYGLKREDQVFEVPCNE
 TQTLNSLPKNVYRVRISAGTGYGVPSQWQHRTPGVHNQSHVFPAPAEKLVRAK
 MESLVSMQPPPHPTQISGKLYMRVETEEADGDRPGGRGDQAMVDGVPRLKKV
 KOYELTLVPLGLKEDFTTVKIVNYTVRFGWLRNLSLVYTTSSGEDILIGLKP
 AESNSTSLIWLKPKPDKFTTVKIVNYTVRFGWLRNLSLVYTTSSGEDILIGLKP
 FTKYFAVGHGVMDMPGFSVVERSTLDRPSTPDLRLSLPTSTVRLHWCPTE
 PNGEIVELILYNNHTQPEHOWTLLTTEGNI FSAEVHGLSDTRYFPFGMARTVGP
 GPSPRLQDVIQKTPSDLSLDVHVTGIIYGVCLGLLCLLACHCAGLRSSHREALPG
 LSSSGTPGNFALYTRGPPSPAAHELSLVHPRPQDMSPPPDVEDKAVHSLMG
 GSVDCRSHSKRIINWAGQGNWAGCPLQSGRPRPALTRALPPAGTGTL
 LQALVYVADLQSGADWLQRELGCGQPTTSGPERLTCLPEAASASCSDLPSPSTAE
 EAPGSKCPKALCPITVSPSLPRAPVSSAQV"

BASE COUNT 1362 a 1770 c 1669 g 1500 t
 ORIGIN

Alignment Scores:

Pred. No.: 66-250 Length: 6301
 Score: 6351.50 Matches: 1220
 Percent Similarity: 95.40% Conservative: 4
 Best Local Similarity: 95.09% Mismatches: 8
 Query Match: 96.18% Indels: 52
 DB: 10 Gaps: 2

US-09-754-997A-2 (1-1252) x AK122535 (1-6301)

QY 21 ArgGlyGluLeuProLeuProGlnGluThrThrVallyLeuSerCysAspGluGlyPro 40
 DB 39 AGAGGGAGCTGGCCATTGCCCCAGAGAGACACTGTCAAGCTGAGCTGTGATGAGGACCC 98
 QY 41 LeuGlnValIleLeuGlyProGluGlnAlaValValLeuLeuAspCysThrLeuGlyAlaThr 60
 DB 99 CTGCAGTGTATCTCTGGGCCCTGAGCAGGCTGTGTGTCTGTGACTGCACTTGGGGCTACA 158
 QY 61 AlaAlaGlyProProThrArgValThrTrpSerIysAspGlyAspThrValLeuGluHis 80
 DB 159 GCTGTGGGCCCTCCAGCAGGGTGACATGGAGCAGGATGAGACACTGTACTAGAGCAT 218
 QY 81 GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlnGlu 100
 DB 219 GAGAACCTTGCACTGCTACCCCAATGGCTCCCTGTGGCTGTCTCAACCTTAGAGGAGNA 278
 QY 101 AspSerAspGluGluAlaLeuArgIleTrpIysValThrGluGlySerTrpSerCys 120
 DB 279 GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTGT 338
 QY 121 LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValIysLeuAlaThr 140
 DB 339 CTGGCCCAACAGCCCGCTAGGAGTGTGGCCAGCCAGGTGTCTGTGTGTCAGCTTCCACA 398
 QY 141 LeuGluAspPheSerLeuHisProGluSerGlnIleValGluAsnGlyThrAlaArg 160
 DB 399 CTCGAAGACTTCTCTCTGCACCCCGAGTCCCAAGATGTGGAGGAGACGGGACGACGC 458
 QY 161 PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnVal 180
 DB 459 TTTGAATGCCACACCAAGGGCTTCCAGCCCCCATCATTTACTTGGGAAAGACCCAGGTG 518
 QY 181 ThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuGlnIleLeuAsp 200
 DB 519 ACCGTGCTGAGGAGCTCCGGCTCATCACTCTTCCCAATGGGCTCTCCAGATCTTAGAT 578
 QY 201 ValGlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArg 220
 DB 579 GTCCAGGACAGTATGAGGCTCTTACCCTGCTGGGCCACCAATTCAGCCCGCCCAACGA 638
 QY 221 PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly 240
 DB 639 TTCACCCAGGAGGCTCTCTCTCTGTGGCCCTCAGAGGGTCTTTGGAGGCTACACAGGGG 698
 QY 241 GlnAspValIleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnVal 260
 DB 699 CAGGATGTGGTCAATTGTGGCAGCCCAAGAACACACCGTAGTGTCTGGACAGAGTGA 758
 QY 261 ValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGln 279
 DB 759 GTGATGGAGTGGTGGCTCTGTCTGACCCCAACCCCTTTTGTGTCTTGGGTCCGACA -GGG 817
 QY 279 ----- 279
 DB 818 GGATCAGAGTAAGAGGAACCTGGGGATTCAAACACAGCGCTGTGTATGTCAGCTTTGCAGGG 877
 QY 279 ----- 279
 DB 878 ATGTAGTGGGAACCTCGTGTAGTTGGAGAGGTGTAGAGGGGACAGTTCACAGCGTCCCCCTC 937
 QY 280 -----AspGlyLysProIleSerThrAspValIleValLeu 291
 DB 938 CAGCCCTGCTCTTCTCTGTCTACCCAGATGGAAAGCCATATCTCCACGGATGTATCGTCTG 997
 QY 292 GlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHisSerGlyValTrpValCys 311
 DB 998 GGCCCGACCAATCTACTCATGCCAGCGCGAGCTCTGGACTCTGGAGTCTATGTGTGC 1057
 QY 312 ArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeu 331

Db	1058	CGAGCCAAACAGCCCGCAGCGGTGATTTGCGCCACTGCGGCTGCTGAGCTCCGAGTGCTT	1117
Qy	332	AlaalaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAla	351
Db	1118	GCTCCCGCAGCATCTCGCAGCGCCCGAGCGCTCTCGCGAGCGCGGCGCAGACCGCG	1177
Qy	352	ArgPheValCysArgAlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGly	371
Db	1178	CGCTTCGTGTGCGGGGAGTCCGGGAGCCACGCGCCCGCTGCACCTGGCTGCACGCGGG	1237
Qy	372	IleProLeuArgProAsn-GlyArgValIysValGlnGlyGlyGlySerLeuValII	391
Db	1238	ATCCCGTTGGACCAATGGCGCGCTCAAGGTGCGAGGGGTGGCGGCGAGCTTGGTCA	1297
Qy	391	eThrGlnIleGlyLeuGlnAspAlaGlyTyTrpGlnCysValAlaGluAsnSerAlaG1	411
Db	1298	CACCTCAGATGGCCCTTCAGGACGCTGGCTACTACAGTGGGTAGCAGAAACAGCGCGGG	1357
Qy	411	YThrAlaCysAlaAlaProLeuAlaValValAlaGluGlyLeuProSerAlaPr	431
Db	1358	AACCTGCTGTGCGCTCGGCCCTCGCGGTAGTGTGGCGAGGGGCTGCCAGCGGCC	1417
Qy	431	oThrArgValThrAlaThrProLeuSerSerSerSerValIleValAlaTrpGluArgPr	451
Db	1418	GACTCGGCTCACGCCACGCGCTGAGCAGCTCTCTGTGTGTGGCTGGGAGCGGCC	1477
Qy	451	oGluLeuHisSerGlnIleIleGlyPheSerLeuHisTyGlnIysAlaArgGlyVa	471
Db	1478	TGAGTTGCACAGCAGCAAAATCATTTGGCTTCTCTCTACCTACCAAAAGGCAAGGGAGT	1537
Qy	471	lAspAsnValGluTyTrpGlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAs	491
Db	1538	GGACNATGTGAGTACAGATTGTCAGTTAAACATGACACACAGAGCTGCAGGTTCCGGA	1597
Qy	491	pLeuGluProAsnThrAspTyTrpGluPheTyValValAlaTySerGlnLeuGlyAlaSe	511
Db	1598	CCTGAAACCCCAACAGCATATGAGTTCTACGTGTGGCTTACTCCAGCTGGGGCCAG	1657
Qy	511	rArgThrSerSerProAlaLeuValHisThrLeuAspAspValProSerAlaAlaProG1	531
Db	1658	CCGAACCTCCAGCCCGCCCTGGTGATACACTGGAGATGTCCCGCAGCGAGACCCCA	1717
Qy	531	nLeuThrLeuSerSerProAsnProSerAspIleArgValAlaAlaTrpLeuProLeuProSe	551
Db	1718	GCTTACCTTGTCCAGCCCAACCCCTCGGACATCAGGGTGGCATGGCTGCCCTGCCCTC	1777
Qy	551	rSerLeuSerAsnGlyGlnValLeuLysTyLysIleGluTyTrpGlyLeuGlyLys---	570
Db	1778	CAGCTGAGCAATGGACAGAGTGCTGAAGTACAAGATAGAGTACCGTTTGGGGAAGGA	1837
Qy	570	uAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLe	590
Db	1838	AGATCAGGTTTCTCCACCGAGGTGCTGGAAATGAGACACAACTTACGTTAAACTCACT	1897
Qy	590	uGlnProAsnLysValTyArgValArgIleSerAlaGlyThrGlyAlaGlyTyTrpGlyVa	610
Db	1898	TCAGCCAAACAAAGTGTACCGAGTCCGGAATTCAGCTGGCACTGGCGCTGGCTATGAGT	1957
Qy	610	lProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGlnSerHisValProPh	630
Db	1958	CCCTTCTCAGTGGATGACAGCAGACACCTGGTGTGCACAAACAGAGCCATGTTCCTT	2017
Qy	630	eAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuValValSerTrpGlnPr	650
Db	2018	TGCCCCCTGCAGAAATGAAGGTGAGGGCAAGATGGAGTCCCTGGTGGTGTCAATGGCAGCC	2077
Qy	650	oProProHisProThrGlnIleSerGlyTyTrpLysLeuTyTrpGlyGluValGlyThrG1	670
Db	2078	GCCCCCTCACCCACCCAGATCTCTGGATACAAACTCTACTTGGAGAGGTGGGAAACGA	2137
Qy	670	uGluAlaAlaAspGlyAspArgProProGlyGlyArgGlyAspGlnAlaTrpAspValG1	690
Db	2138	GGAGGAGGCAGATGGTGAACCCGCCCGCCAGGGGGTGTGGAGATCAAGCTTGGGACGTGG	2197
Qy	690	yProValArgLeuLysLysValIysGlnTyTrpGluLeuThrGlnLeuValProGlyAr	710
Db	2198	GCCCGTGGCGCTGAAGAGAAAGTGAAGCAGATATGAACCTGACCCAGTTAGTCCCTGGCAG	2257
Qy	710	gProTyTrpGluValLysLeuValAlaPheAsnLysHisGluAspGlyTyTrpAlaAlaValTr	730
Db	2258	GCTGTACGAGGTGAAGCTCGTAGCTTCAACAAACACGAGGACGGCTACGCTGCTGTGTG	2317
Qy	730	plysGlyLysThrGluLysAlaProThrProAspLeuProIleGlnArgGlyProProLe	750
Db	2318	GAAGGCAAGACGGAGAGGCCGCCACGCCAGACCTGCTATCCAGAGGGGCCACCGCT	2377
Qy	750	uProProAlaHisValHisAlaGluSerAsnSerSerThrSerIleTrpLeuArgTrpLy	770
Db	2378	GCCTCTGCCCATGTCCACGACAGTCAAAACAGCTCCACTTCCATTTGGCTTGGTGGAA	2437
Qy	770	sLysProAspPheThrThrValLysIleValAsnTyThrValArgPheGlyProTrpG1	790
Db	2438	GAAGCCAGACTTTTACCACTGTCAAGATTGTCAACTACACTGTACGCTTCGGGCCCTGGGG	2497
Qy	790	yLeuArgAsnAlaSerLeuValTyTrpTyThrSerSerGlyGluAspIleLeuIleG1	810
Db	2498	GCTCAGGAATGCTTCCCTGCTCACCTATATACCACTCTGGAGAGACATTTCTCATTTGG	2557
Qy	810	yGlyLeuLysProPheThrLysTyTrpGluPheAlaValGlnSerHisGlyValAspMetAs	830
Db	2558	CGGCTCGAAACCATTTACCACTACGAGTTTGGGTGTACAGTCCCAACGAGTGGATATGA	2617
Qy	830	pGlyProPheGlySerValValGluArgSerThrLeuProAspArgProSerThrProPr	850
Db	2618	TGGGCGCTTGGCTCGCTCGTAGAACCGCTCCACCCCTCGCTGACCGGCTTCAACACCTCC	2677
Qy	850	oSerAspLeuArgLeuSerProLeuThrProSerThrValArgLeuHisTrpCysProPr	870
Db	2678	TTCTGACCTGGCGCTGAGCCCTGACACCATCCACCGTTCGGGTATACCTGGGTGTCCTCC	2737
Qy	870	oThrGluProAsnGlyGluIleValGluTyTrpLeuLysLeuTyTrpSerAsnAsnHisThrG1	890
Db	2738	CACGAGCCCAATGGTGAGATTGTGGAGTATCTATTTCTTACAGCAACACACACACCCA	2797
Qy	890	nProGluHisGlnTrpThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHi	910
Db	2798	GCCCGAACCACTAGTGGACACTGCTCACCAACAGAGGGAAACATCTTCAGTGCAGAGGTCCA	2857
Qy	910	sGlyLeuGluSerAspThrArgTyPhePheLysMetGlyAlaArgThrGluValGlyPr	930
Db	2858	TGGCCTTAGAGAGTGAACACTCGGTATTCTTCAAGATGGGAGCCCGCACAGAGGTGGGCC	2917
Qy	930	oGlyProPheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSerLe	950
Db	2918	TGGGCGCTTTTCCCGCTTGCAGGATGTGATTACTCTGCAAAAGACATTTCTCAGACTCCTT	2977
Qy	950	uAspValHisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuLe	970
Db	2978	GGATGTGCACCGCGCTCACGGGCATCATCGTGGGTGTCTGCTGGGCGCTTCTCTGCTCCT	3037
Qy	970	uAlaCysMetCysAlaGlyLeuArgGlnSerSerHisArgGluAlaLeuProGlyLeuSe	990
Db	3038	GGCTTCGATGTGTGCTGGCGCTACGACGAAGCTCCCAAGGAAGCCCTGCCCCGATGTGC	3097
Qy	990	rSerSerGlyThrProGlyAsnProAlaLeuTyThrArgAlaArgLeuGlyProProSe	1010
Db	3098	CTCTCAGGCACCCAGGAACCCAGCGCTCTACACAGAGCTCGACTTGGGCCCCCAG	3157
Qy	1010	rValProAlaAlaHisGluLeuGluSerLeuValHisProArgProGlnAspTrpSerPr	1030
Db	3158	TGTCCCTGCTGCCCATAGTTGGAGTCCCTCGTGCACTCTCGTCCCGCAGGATTTGGTCCCC	3217
Qy	1030	oProProSerAspValGluAspLysAlaGluValHisSerLeuMetGlyLysSerValSe	1050
Db	3218	ACCACCTTCAGATGTGGAGACAAAGCTGAAGTACACAGCCTTATGGTGGCAGTGTTC	3277


```

QY 1050 rAapCyArgGlyHisSerLysArgLysIleSerTrpAlaGlnAlaGlyGlyProAsnTr 1070
DB 3278 AGATTGCGCGGCGCCACTCAAGAGAAAGATCTCTCGGCTCAGGCGGCGGCAAACTG 3337
QY 1070 pAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuTh 1090
DB 3338 GCGAGCTCTCGGCGAGCTGTGAGCTGCCCGAGGTAGTGTCCAAAGCGCGCTCTGAC 3397
QY 1090 rArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeuLeuLeuGlnAlaLeuValTy 1110
DB 3398 CGGTGCTCTGCTGCTCTCAGCGGGAACCGGCGAGACTGTGCTGCTGCAAGCCCTGGTATA 3457
QY 1110 rAapGlyIleLysSerArgGlyArgLysIleProSerProAlaCysArgAaGlnValGl 1130
DB 3458 TGACGCCATAAGACCAAGCGGAGAAAGACCGTCCCGAGCTCTGAGGATCAGGTGGA 3517
QY 1130 uAlaGluValIleValHisSerAspPheGlyAlaSerLysGlyCysProAspLeuHisLe 1150
DB 3518 AGCTGAGGTCAVTTGTCCACTCGGACTCGGTCGATCAAAAGGATGCTCTGACCTCCACT 3577
QY 1150 uGlnAspLeuGluProGluProLeuThrAlaGluThrLeuProSerThrSerGlyAl 1170
DB 3578 CCAAGACTGAGCGCAGAGGACCACTGACTGCGAGAGACTCTGCCCTTCAAGTCTGGAGC 3637
QY 1170 aValAspLeuSerGlnGlyAlaAspTrpLeuGlyArgGluLeuGlyCysGlnProTh 1190
DB 3638 TGTGATCTGTCTCAAGAGAGCAGACTGCTGCGGAGGAGCTGCGGAGGCTGCCAACCA 3697
QY 1190 rThrSerGlyProGluArgLeuThrCysLeuProGluAlaAlaSerAlaSerCysCy 1210
DB 3698 AACCAAGTGGCGCAGAGAGCTCACCTGCTGTCAGAGAGCAGTGCCTCTCTGCTCTG 3757
QY 1210 sSerAspLeuGlnProSerThrAlaIleGluAlaProGlyLysSerCysGlnProLy 1230
DB 3758 CTCAGACTCCAGCCAGCAGACTGCTGTATAGAGAGGCCCTCGGAAAGCTCCAGGCCCA 3817
QY 1230 sAlaLeuCysProLeuThrValSerProSerLeuProArgAlaProValSerSerAlaGl 1250
DB 3818 AGCCCTGTGCTCTTAAGCAGTCAGCCAGCCAGCCCTCCAGGCGCCCTGCTCTCTGCTCA 3877
QY 1250 nValPro 1252
DB 3878 GGTCCCC 3884

RESULT 7
AB052622 AB052622 6485 bp mRNA linear PRI 16-MAR-2002
LOCUS Homo sapiens hDDM36 mRNA, complete cds.
DEFINITION AB052622
ACCESSION AB052622
VERSION AB052622.1 GI:19570397
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
MURAKAMI,H., NAKAMATA,T., OKA,M., KIYONO,T., SASAKI,M.S., NAKAMURA,T.,
AOYAMA,T., NEGAYAMA,S., YAMAMOTO,H., HOSAKA,T.,
Up-regulation of a ras effector and down-regulation of a cell
adhesion molecule are associated with transformation of osteoblasts
Unpublished
JOURNAL 2 (bases 1 to 6485)
REFERENCE Toguchida,J., Nakamata,T., Murakami,H., Nakayama,T. and Nakamura,T.
AUTHORS Direct Submission
TITLE Submitted (11-DEC-2000) Junya Toguchida, Kyoto University,
JOURNAL Institute for Frontier Medical Sciences; 53 Kawahara-cho, Shogoin,
Sakyo-ku, Kyoto, Kyoto 606-8507, Japan
(E-mail:togjun@frontier.kyoto-u.ac.jp, Tel:81-75-751-4134,
Fax:81-75-751-4144)
FEATURES
Location/Qualifiers
1..6485
/organism="Homo sapiens"

```

```

/mb_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/clone="hDDM36"
/cell_line="NH08"
/cell_type="osteoblast"
/notes="similar to mouse DDM36 and Nope
(NM_020043)-vector:phDDM36"
1..6485
/gene="hDDM36"
211..3963
/gene="hDDM36"
/notes="similar to mouse DDM36 and Nope (NP_064427)"
/codon_start=1
/evidence="experimental"
/product="hDDM36"
/protein_id="BAB86306.1"
/db_xref="GI:19570398"
/translat="MARGDAGRGRLALFCLLAARGELLFQETTVELSCGVGPIQ
VILGPEQAAVLNGLSAGAAAGPPTVTMSKQDITLLEHDLHLLPGLSLWSOPFAPN
GSDSPVEAVGVIKGYSLAAGPLVGLAVKLAATLADFLSLHPESOTVBERNGTA
RFRCHLEGLPAPIITWEKQVTLPEEPRLVLNGLQIILDQESDAGPFRVATNSA
RQHSBALLSVAHRSGLASTRGQDVVIVAAPENTTVVQSVVMBVCASADPTPVFS
WVRQDKPISDVIIVLGRITNLLIANAQFVHSVYVCRANKPRTDPAFAAEFLVLA
PAITQAPALSRTRASTARFVCRASGEPRPALRWLHNGAPLRNGRVKVOGGGSLVI
TOIGLDAGTYQCVAENAGMACAAALAVVREGLPSPATRTVATPLSSAGVLAWE
RPMHSQIIGFSLHYOKARGMDNVQPAVNDTTELOVRDLEPNTDYEFVYVAYSO
LGSRTSTPALVHTLDDVPSAQLSLSPNSDIRVAMLPPLPSLNGQVVKYKIEY
GLGKEDIFSTEVRGNETQMLNSLQPNKYRVRISAGTAPGAPGAPGAPGAPGAPGAPG
RQSHVPPAPAEKVQAKMSLVWSQPPHPTQISGKLYRVEYGAEEANGDLPGG
NQDQADVGVRLKKYKQYELTQVLPGRPLYKLVAFNKHEDGYAAVWKGDKLAPA
PPIORGPPLPPAHVHAESENSTSIWRWKKPDPPTVKIIVNYTVRSPWGLRNASLV
TYTSSGDIIGLGLKPTKYEPAVQSGVMDGPGVSVVERSTLPDRBSTPPSDLRL
SPLTPSVRLHWCPTPEPNGEIVEYLYLYSSNTQPEHQTLLTQGNISAEVHGLE
SDTRYFPMGARTVSGPFPFRLQDVITLQKLKSDLSDMHSTVGIIVGVCLLCLLA
CMAGLRSRPHRESLPGLSSTATFNPALYSRRLPSPPPAAHELSLHVHPHQLP
PTEVEDRAEVHSLGGVSEGRSHSKRIKISWAPCSGLSWAGCSLWAGCPQAGPRA
LTRALLPAGTGTGTLIQAUVYDAIKNGRKTKPPACRNOVEARVIVHSPFASNGNP
DLHLQLEPDPPLPEAPDLISGVDPQCAWLDRELGCLELAAPGPDRLTCLPEAA
SASCSVPDLQGVLESTPQDSQLKSPCLGASPLPRFSPVSSA"
BASE COUNT 1293 a 1926 c 1772 g 1494 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 2,87e-225 Length: 6485
Score: 5746.00 Matches: 1098
Percent Similarity: 92.09% Conservative: 54
Best Local Similarity: 87.77% Mismatches: 96
Query Match: 87.01% Indels: 5
DB: 9 Gaps: 2

```

US-09-754-997A-2 (1-1252) x AB052622 (1-6485)

```

QY 1 MetAlaArgAlaAspThr-----GlyArgGlyLeuLeuValLeuThrPheCysLeuLeu 18
DB 211 ATGCGCGGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 270
QY 19 SerAlaArgGlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCysAspGlu 38
DB 271 GCCGCGCGCGGGAGCTGTGTTGCCCGCAGAGACGACTGTGGAGCTGAGCTGTGGAGTG 330
QY 39 GlyProLeuGlnValIleLeuGlyProGluGlnAlaValLeuLeuAspCysThrLeuGly 58
DB 331 GGGCCACTGCAAGGATCTCGGCGCCAGAGAGCGCTGACGTCTAACTGTAGCTGTGGG 390
QY 59 AlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeu 78
DB 391 GCTGCTGCCCTGGACCCCCCACCAGCGGTGACCTGGAGCAAGATGGGACACCTGCTG 450
QY 79 GluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlu 98
DB 451 GAGCAGCACCACTTACACCTGCTGCCCAATGGTTCCCTGTGGCTGTGCCAGCCACTAGCA 510

```


Db 2668 GAGTTTGCAGTGCAGTCTCAAGCGCTGGACATGATGGCCCTTTCCGGCTCTGTGGTGAG 2727
 Qy 839 ArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeu 858
 Db 2728 CGCTCCAGCTCGCTGACGGGCGCTCCACACCCCATCCGACCTCGAGCTGAGCCCGCTG 2787
 Qy 859 ThrProSerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluVal 878
 Db 2788 ACACCGTCCAGGTTCCGCTGCACTGGTGGCCCCCAGAGCCCAACGGGGAGATCGT 2847
 Qy 879 GluTrpLeuLeuLeuSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeu 898
 Db 2848 GAGTATCTGATCTGTACAGCAGCAACACACGACGCTGACCACTGAGACCTTGGCTC 2907
 Qy 899 ThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyr 918
 Db 2908 ACCACGACGAGGAACATCTTCAGTCTCGAGTCCATGCGCTGGAGAGCGACACTCGGTAC 2967
 Qy 919 PhePheLeuMetGlyValArgThrGluValGlyProGlyProPheSerArgLeuGlnAsp 938
 Db 2968 TTCTTCAGATGGGGCGCGCACAGAGTGGGACCTGGGCGCTTTCTCCGCGCTGCAGGAT 3027
 Qy 939 ValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIle 958
 Db 3028 GTGATCAGCTCCAGAGAGAGCTGTACACTCGCTGGACATGCACTCAGTCACGGGCATC 3087
 Qy 959 IleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArg 978
 Db 3088 ATCGTGGGTGTCTGCTGGGCTCTCTGGCTCTGCGCTGCACTGTGTGTGGCTGGCG 3147
 Qy 979 GlnSerSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnPro 998
 Db 3148 CGCAGCCCCACAGGGAAATCCCTCCAGGCTGTCTCCACCGCCACCCCGGGGAATCCC 3207
 Qy 999 AlaLeuTrpThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGlu 1018
 Db 3208 GCGCTGTCTCAGAGCTCGGCTGGCCCCCCCCAGCCCCCGAGCTGCGCATGAATTGGAG 3267
 Qy 1019 SerLeuValHisProArgProGlnAspTrpSerProProSerAspValGluAspLys 1038
 Db 3268 TCCCTTGTGCACCCCATCCCGAGGACTGTGTCGCCGCCACCTCAGAGCTGGAGGACAGG 3327
 Qy 1039 AlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArg 1058
 Db 3328 GCTGAAGTGCACAGCTTATGGGTGGCGGTGTTCTGAAGGCGGAGTCACTCCAAAGA 3387
 Qy 1059 LysIleSerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGlu 1078
 Db 3388 AAGATCTCTGGGCTCAACCAAGCGGCTGAGCTGGGCTGTCTCGGCGAGGCTGTGAG 3447
 Qy 1079 LeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGly 1098
 Db 3448 CTGCCCCAG---GCAAGGCCCGCGCGGCTCTGACCCCGGCGCTGCTGCCCGCTGCTGGA 3504
 Qy 1099 ThrGlyGlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArg 1118
 Db 3504 ACTGGGCGACGCTGTCTGAGGCTCTGGGTGTACGACGCCATAAAGCGCAATGGGAGG 3564
 Qy 1119 LysLysProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAsp 1138
 Db 3565 AAGAAAGTCAACCCCGAGCTGCAAGAACAGGTTGGAGGCTGAAGTCAATTGTCCACTCTGAC 3624
 Qy 1139 PheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluPro 1158
 Db 3625 TTTAGTGCATCTAACCGGAACCTGACCTCCATCTCCAGAGCTTGGAGCTGAGAGCCCC 3684
 Qy 1159 LeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAsp 1178
 Db 3685 CTGCTCCAGAGCTCTGATCTCATCTCGGGTGTGGGATCCAGGCGAGGGGCGACCC 3744
 Qy 1179 TrpLeuGlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThr 1198
 Db 3745 TGGCTGGACAGGAGTTGGAGGGGTGTGAGCTGGCGAGCCCCCGGCGCACAGACTTACC 3804

Qy 1199 CysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAla 1218
 Db 3805 TGCTTGCAGAGGCGAGCGAGTCTTCTGCTCTTACCGGACCTCCAGCCAGGCGAGTG 3864
 Qy 1219 IleGluGluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSer 1238
 Db 3865 CTAGAGAGACCCCTGGAGATAGCTGCCAGTCAAAATCCCCCTGCTCCTAGGAGCCAGC 3924
 Qy 1239 ProSerLeuProArgAlaProValSerSerAla 1249
 Db 3925 CCAGGCTGCCAGATCCCGGCTCTCTCTCT 3957
 RESULT 8
 AX418406 3741 bp DNA linear PAT 18-JUN-2002
 LOCUS Sequence 1 from Patent WO0206329.
 DEFINITION AX418406
 ACCESSION AX418406
 VERSION AX418406.1 GI:21523336
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 1 unclassified.
 REFERENCE
 AUTHORS Rastelli, L., Shimkets, R.A., Zerhusen, B., Malyanar, U.M. and Padigaru, M.
 TITLE Human polynucleotides and polypeptides encoded thereby
 JOURNAL Patent: WO 0206329-A 1 24-JAN-2002;
 Curagen Corporation (US)
 FEATURES
 source 1. 3741
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 /note="NOVX nucleic acid with homology to NOPE-like protein"
 BASE COUNT 708 a 1209 c 1150 g 673 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,38e-220 Length: 3741
 Score: 5610.00 Matches: 1086
 Percent Similarity: 91.07% Conservative: 56
 Best Local Similarity: 86.60% Mismatches: 101
 Query Match: 84.95% Indels: 13
 DB: 6 Gaps: 8
 US-09-754-997A-2 (1-1252) x AX418406 (1-3741):
 Qy 1 MetAlaArgAlaAspThr-----GlyArgGlyLeuValLeuThrPheCysLeuLeu 18
 Db 1 ATGGCGGGGGGAGCGCGCGCGCGGCTCTCGGTTGACCTTCTGCTGTG 60
 Qy 19 SerAlaArgGlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCysAspGlu 38
 Db 61 GCCCGCGCGGGAGCTGTGTTGCCCGAGAGACGACTGGAGCTGAGCTGGAGTG 120
 Qy 39 GlyProLeuGlnValIleLeuGlyProGluGlnAlaValLeuLeuAspCysThrLeuGly 58
 Db 121 GGGCCACTGCAAGTGATCTCGGGCCAGAGCGGCTGAGTGTAACTGTAGCTGGG 180
 Qy 59 AlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeu 78
 Db 181 GCTGCTGCGCTGAGACCCCGCCACCGAGGTGACCTGGAGCAAGATGGGACACCTGCTG 240
 Qy 79 GluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlu 98
 Db 241 GAGCACGACCACTTACACCTGCTGGCCAAATGTTCCCTGTGGCTGTCCAGCCACTAGCA 300
 Qy 99 GlnGluAspSerAspAspGluAlaLeuArgIleTrpLysValThrGluGlySerTyr 118
 Db 301 CCCAATGGCAGTGCAGT--CAGTCCCTGAGGCTGTGGGG-GTCATTGAAGGCACTAT 357

QY	62	AlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGluHisGlu	81	1201	GTGGTGGCGAGGGGCTGCCAGCGCCCGACTCGGGTCAACAGCCAGCGCGCTGAGCAGC	1260
DB	121	GCTGGGCGCTCCGACAGGCTGACATGAGCAGCATGAGACACTGTACTAGACATGAG	180			
QY	82	AsnLeuHisLeuLeuProLanGlySerLeuTrpLeuSerSerProLeuGluGlnAsp	101	442	SerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleLeuGlyPhe	461
DB	181	AACCTGCACCTGTACCAATGGCTCCCTGTGGCTGTCTCTCAACCCCTAGAGCAAGAAGAC	240	1261	TCCTCTGTGCTGGTGGCGCTGGAGCGGCTGAGTTGCACAGCGAGCAATCATTTGGCTTC	1320
QY	102	SerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTrpSerCysLeu	121	462	SerLeuHisTrpGlnLysAlaArgGlyValAspAsnValGluTrpGlnPheAlaValAsn	481
DB	241	AGCGATGATGAGGAAGCTCTTAGATCTGGAAGTCACTAGAGGCGAGCTATTCTCTGTCTG	300	1321	TCCTCTCACTACCAAAAGGCAAGGAGTGGACATGTGGAGTACCAAGTTTGCAGTAAAC	1380
QY	122	AlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThrLeu	141	482	AsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPheTrp	501
DB	301	GCCCCACAGCCCGCTAGGAGTGGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGCCACACTC	360	1381	AATGACACACAGAGCTGCAGGTTCCGGACCTCGGAACCCCAACACGAGTATTGAGTTCTAC	1440
QY	142	GluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArgPhe	161	502	ValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThr	521
DB	361	GAAGACTTCTCTGCAACCCGAGTCCAGATTTGTGAGAGAGAAACGGGACAGCAGCTTT	420	1441	GTGGTGGGCTACTCCAGCTGGGGGCGAGCGCAACTCCAGCCAGCCCTGGTGATACA	1500
QY	162	GluCysHisThrLysGlyLeuProAlaProIleThrTrpGluLysAspGlnValThr	181	522	LeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAsp	541
DB	421	GAATGCCACACCAAGGGGCTTCAGGCCCCCATCAATTCTTGGAAAGGACCAAGTGACC	480	1501	CTGGAGCATGTCCCGAGCGCAGCACCCAGCTTACTTGTCCAGCCCAACCCCTCGGAC	1560
QY	182	ValProGluGluProArgLeuLeuThrLeuProLysTrpLeuLeuGlnIleLeuAspVal	201	542	IleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTrp	561
DB	481	GTGCTGAGGAGGAGCCCGGCTCATCTCTCCCAAGTGGCTCTCCAGATCTAGATGTC	540	1561	ATCAGGGTGGCATGCTGCCCTGCCCTCCAGCTGAGCAATGGACAGGTGCTGAAGTAC	1620
QY	202	GlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArgPhe	221	562	LysIleGluTrpGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsn	581
DB	541	CAGGACAGTGATGACAGGCTCTACCGCTGGTGGCCACCAATTCAGCCGCGCAACGATTC	600	1621	AAGATAGAGTACGGTTTGGGGAAGAGATCAGGTTTCTCCAGCCAGGTGCTGGAAAT	1680
QY	222	SerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGln	241	582	GluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTrpArgValArgIleSer	601
DB	601	AGCCAGGAGGCGCTCGTCACTGTGGCCCTCAGAGGCTCTTTGAGGCTACCAAGGGGCGAG	660	1681	GAGACACAACCTTACGTTAACTCACTTCAGCCAAACAAGTGTACCGAGTCCGAGTTTCA	1740
QY	242	AspValValIleValAlaProGluAsnThrThrValValSerGlyGlnAsnValVal	261	602	AlaGlyThrGlyValaGlyTrpGlyValProSerGlnTrpMetGlnHisArgThrProGly	621
DB	661	GATGTGTCATTTGGCAGCCCGCAGAGAACCAACCGTAGTGTCTGACAGAAATGTAGTG	720	1741	GCTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGGATGCAGCACAGGACACCTGGT	1800
QY	262	MetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAspGly	281	622	ValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMet	641
DB	721	ATGGAGTGGCTGGCTCTGTGACCCCGCCCTTTGTGCTCGGTGGGTCGACAGGATGA	780	1801	GTGCACAACAGAGCCATGTTCCTTTGCCCCCTGCAGAAATGAAGGTGAGGGCAAGATG	1860
QY	282	LysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAla	301	642	GluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTrpLys	661
DB	781	AAGCTATCTCCACGATGTATCGTTCTGGGCGGACCAATCTACTCATCGCAGCGCG	840	1861	GAGTCCCTGGTGGTGTATGGCAGCGCCCGCTCCACCCACAGATCTCTGATACAA	1920
QY	302	GlnProArgHisSerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAspPhe	321	662	LeuTrpTrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProProGlyGly	681
DB	841	CAGCCTCGGCACCTCTGAGTCTATGTCTGCGAGGCCAACAGCCCTCTACCGGTGACTTC	900	1921	CTCTACTGGGAGAGTGGGAACAGAGGAGGAGGAGATGGTGACCGCCCCCAGGGGT	1980
QY	322	AlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGlu	341	682	ArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysLysValLysGlnTrp	701
DB	901	GCCACTGGCGCTGCTGAGCTCCGAGTCTGCTGCTGCCACGCCATCTCGCAGGACCCCGAG	960	1981	CGTGGAGATCAAGCTTGGACGCTCGGGCGCGTGGCGCTGAGAGTGAAGAGTAT	2040
QY	342	AlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluPro	361	702	GluLeuThrGlnLeuValProGlyArgProTrpGluValLysLeuValAlaPheAsnLys	721
DB	961	GCGCTCTCGCGGACCGCGGCGCAGCACCGCGCTTCTGTGCGCGGCTCGCGGAGCCA	1020	2041	GAACTGACCACTGTAGTCCCTGGCAGCGCTGACAGGTGAAGCTCTTCAACAA	2100
QY	362	ArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLys	381	722	HisGluAspGlyTrpAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAsp	741
DB	1021	CGGCCCGCGCTGCACTGGCTGACAGCGGATCCCGTTGCGACCCCAATGGCGCGCTCAAG	1080	2101	CACGAGACGGCTACGCTGTGTGGAAAGGCGACAGCGAGGAGGCGCCAGCCAGAC	2160
QY	382	ValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTrp	401	742	LeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsnSer	761
DB	1081	GTGACGGCGGCTGGCGGACGTTGGTCACTCATCATCGATCGGCTCGCAGGAGCTGGCTAC	1140	2161	CTGCCCTATCCAGAGGGGCGCACCGCTGCTCTCTGCCCATGTCCACCGAGGTCAACAGC	2220
QY	402	TyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaVal	421	762	SerThrSerIleTrpLeuArgTrpLysProAspPheThrThrValLysIleValAsn	781
DB	1141	TACCAGTGGCTAGCAGAAACAGCGCGGAACCTGCCGTGCGCGCTGCGCGCTGCGCGTA	1200	2221	TCCACTTCATTTGGCTTCGGTGGAGAGCCAGACTTTTACCCTGTCAGATTGTCAAC	2280
QY	422	ValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSer	441	782	TyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTrpTrpThr	801

Db 2281 TACACTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCTCGTCACTACTATACC 2340
 Qy 802 SerSerGlyGluAspIleLeuIleGlyGlyLeuLysProPheThrLysThrGluPheAla 821
 Db 2341 AGCTCTGGAGAGACATCTCATTTGGCGCCCTGAAACCATTTTACCAAGTACGAGTTTGG 2400
 Qy 822 ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr 841
 Db 2401 GTACAGTCCCAACGAGTGGATATGGATGGCCCTTTGGCTCGCTGCTAGAACGCTCCACC 2460
 Qy 842 LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer 861
 Db 2461 CTGCTCTGACCGCCCTTCAACACTCTCTTGACCTGCGCCTGAGCCCTGACACATCC 2520
 Qy 862 ThrValArgLeuHisThrCysProProThrGluProAsnGlyGluIleValGluThrLeu 881
 Db 2521 ACCGTTGCGTTACACTGTTGTCCTCCCGCCGAGCCCAATGTTGAGATTGTGAGATCTA 2580
 Qy 882 IleLeuThrSerAsnAsnHisThrGlnProGluHisGlnThrThrLeuThrThrGlu 901
 Db 2581 ATTCTCTACAGCAACACACACCCAGCCGACACACAGTGGACACTGCTCACCACAGAG 2640
 Qy 902 GlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgThrPhePheLys 921
 Db 2641 GGAAACATCTTCAGTGCAGAGTCCATGCGCTAGAGAGTGACACTCGGTATTTCTTCAG 2700
 Qy 922 MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
 Db 2701 ATGGAGCCCGCACAGAGTGGGCTGCGCCCTTTCCCGCTTGCAGAGATGTATACT 2760
 Qy 942 LeuGlnGluThrPheSerAspSerLeuAspValHis 953
 Db 2761 CTGCAAGAGACATCTCAGACTCTCTGGATGTGCAC 2796
 RESULT 10
 AB046848
 LOCUS AB046848 5549 bp mRNA linear PRI 22-FEB-2001
 DEFINITION Homo sapiens mRNA for KIAA1628 protein, partial cds.
 ACCESSION AB046848
 VERSION AB046848.1 GI:10047332
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (sites)
 Nagase,T., Kikuno,R., Nakayama,M., Hirose,M. and Ohara,O.
 Prediction of the coding sequences of unidentified human genes.
 XVIII. The complete sequences of 100 new cDNA clones from brain
 which code for large proteins in vitro
 DNA Res. 7 (4), 273-281 (2000)
 JOURNAL 20450683
 MEDLINE 10997877
 PUBMED
 REFERENCE
 AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
 TITLE Direct Submission
 JOURNAL Submitted (03-AUG-2000) Otsu City, Kazusa DNA Research Institute,
 Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdna@kazusa.or.jp,
 URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
 Fax:81-438-52-3914)
 FEATURES
 Location/Qualifiers
 1..5549
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="fhl4247"
 /note="vector:pBluescriptII SK plus"
 1..5549
 /gene="KIAA1628"
 <81..3023
 /gene="KIAA1628"
 /note="Start codon is not identified. fhl4247 cDNA clone

for KIAA1628 has a 40-bp different sequence replacing a
 108-bp sequence at the position between 737 and 846, and
 an 88-bp insertion after the position 983 of the sequence
 of KIAA1628, respectively."
 /codon_start=1
 /product="KIAA1628 protein"
 /protein_id="BAB13454.1"
 /db_xref="GI:10047332"
 /translation="MGRKGRSALSRSLRNVQGVKVIIGRGARGGGGAGTGGG
 GRDLSGIPRSASPLAALPAITQAPALSRTRASTARFVCRASGRPRPALWHLNG
 APLRNGRVKVGSGSLVITQIGLDAGYTCVAENSAGMACAAASLAVVVRGLPS
 APTVATPVSASAVLVAWERPEHSDQIIGFSLHYQKARGMDVRYQFVAVNNDTTEL
 QVRDLPTNTYEFYVAYISQIGASRTSTPALVHTLDDVPSAAPQLSSNPNPISRA
 WLPPLPSLNSQVVKYKIEYGLAGKEDQIFSTEVAGNETQLMLNSLPKVKYVRIRISAG
 TAAGFGAPSOVMHRTSPMNQSHVPAEVLKQAKMSELVSVQPPHPTQISGVK
 LYREVCASBEANGDRPLPGGRGQADWDVGVRLAKKQVVELTOLVGRKLYSVRLVAF
 KNGEDGYAVYKNGTEKAPADMPIQRGPPPLPAHVHAEENSSSTIWLKWKPPFTTV
 KIVNTRFPMGLNLSLVITTSRGLIIGLKEPTFKTEFPAVQSHGVQMDGPPGS
 VVRSKTLDPSPSTPDLRLSPFLSTVRLHWCPTFPNGEIVLYLSNHTQPEH
 QMTLLTTQGNIFSAVHGLSDTRYFFMGARTVEVGPFSRLQDTTLAQKLSDSL
 MHSVTGIIVGCLLCLLACMCAGLARSRRBSLPLGLSSTATPGNPLYSRRLGPP
 SPPAAHLSLVHPHDPQWSPDPVDRARVHSLMGGVSBGSHSKRKISWAQPSG
 LSWAGSWAGCELPOAGPRPALTRALLPACTGTGTLILQALVYDAIKNGRKKSPACR
 NOVBARVIVHSDFSASNGNFDLHLQLEPDPFPPEAPDLISGWDGFGGAADREL
 GGCELAAPGPDRLTCLPRAASASCSPDLQPGVLEETPGDSCQLKSPCLGASPLP
 RSPVSSSA"

BASE COUNT 1157 a 1546 c 1496 g 1350 t
 ORIGIN

Alignment Scores:

Pred. No.: 6,04e-169 Length: 5549
 Score: 4361.50 Matches: 822
 Percent Similarity: 92.55% Conservative: 35
 Best Local Similarity: 88.77% Mismatches: 68
 Query Match: 66.04% Indels: 1
 DB: 9 Gaps: 1

US-09-754-997A-2 (1-1252) x AB046848 (1-5549)

Qy 324 AlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeu 343
 Db 243 TCAGCTGCGCCCTCGCGCCCTTAGCGGCTCCGCCATCTCAGGCGCCGCGGCTG 302
 Qy 344 SerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgPro 363
 Db 303 TCGCGGACGGCGGACGACGCGCTTCTGTGTCGCGCGCTCGGGGAGCCGCGGCA 362
 Qy 364 AlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAlaGlyArgValLysValGln 383
 Db 363 GCGCTGCGCTGGCTGCACACGCGGCGCGCTGCGGCCAACGCGGCGCTCAAGTCCAG 422
 Qy 384 GlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTrpGln 403
 Db 423 GCGCGCGGTGGCAGCTGGTGCATCACAGATCGGCGCTGCGAGGACGCGGCTACTACAG 482
 Qy 404 CysValAlaGluAanSerAlaGlyThrAlaCysAlaAlaAlaProLeuAlaValVal 423
 Db 483 TGCCTGCTGAGAACACGCGCGGAAATGCGTGCCTGCGCGCTGCGCGCTGCGTGTGTG 542
 Qy 424 ArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerSer 443
 Db 543 GCGGAGGGCTGCCAGCGCCCGCCAGCGGGTCACTGCTACGCCACTGAGCAGCTCCGCT 602
 Qy 444 ValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeu 463
 Db 603 GTGTGTGTGCTGGGAGCGCCGAGATGCACAGCAGCAGATCATCGGCTTCTCTCTC 662
 Qy 464 HistTrpGlnLysAlaArgGlyValAspAsnValGluThrGlnPheAlaValAsnAsp 483
 Db 663 CACTACAGAGGCGCGGCGATGGCAATGTGGAATACCAAGTTTGCAGTGAACACGAC 722
 Qy 484 ThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValVal 503

Db	723	ACCACAGAACTACAGGTTCCGGACCTGGAAACCAACACAGATTATGAGTTCTACGTGGTG	782
Qy	504	AlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrIleuAsp	523
Db	783	GCTACTCCACAGCTGGAGCCAGCCGACCTCCACCCACAGACTGGTGACACACTGGAT	842
Qy	524	AspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArg	543
Db	843	GATGTCCCAAGTCAGACACCCAGCTCTCCCTGTCCAGCCCAACCTTCGGACATCAGG	902
Qy	544	ValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIle	563
Db	903	GTGGCGTGGTGGCCCTGCCCCAGCTGAGCAATGGGAGGTGGTGAAGTACAAGATA	962
Qy	564	GluTyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThr	583
Db	963	GAATACGGTTTGGAAAGGAAGATCAGATTCTCTACTAGGTGGGAGGAATAGACA	1022
Qy	584	GlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGly	603
Db	1023	CAGCTTATGCTGAACCTGCTTCAGCCAAACAAGGTGTATCGAGTACGGATTTCGGCTGGT	1082
Qy	604	ThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHis	623
Db	1083	ACAGCAGCCGGCTCGGGGGCCCTCCAGTGGATGCATCACAGGACGCCCACTATGCAC	1142
Qy	624	AsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSer	643
Db	1143	AACAGAGCCATGTGCCCTTTGGCCCTGCAGAGTTGAAGGTGCAGGCAAGATGGAGTCC	1202
Qy	644	LeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyr	663
Db	1203	CTGGTCTGTGATGGCAGCACACCCCTCACCCACACAGATCTCTGGCTACAACTATAT	1262
Qy	664	TrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProProGlyGlyArgGly	683
Db	1263	TGGCGGAGGTGGGGGCTGAGGAGGAGGCCAATGGCGATCGCTGCCAGGGGGCCGTGA	1322
Qy	684	AspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyrGluLeu	703
Db	1323	GACGAGCTTGGGATGGGGGCTGTCCGGCTCAAGAAGAAAGTGAAGCAGTATGAGCTG	1382
Qy	704	ThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGlu	723
Db	1383	ACCAGCTAGTCCCTGGCCGGCTGTACAGGTGAGCTCGTGGCTTCAACAACTAG	1442
Qy	724	AspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuPro	743
Db	1443	GATGGCTATGCAGCAGTGTGGAAGGCAAGACGAGAGGCGCGCCGACCAACATGCCT	1502
Qy	744	IleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsnSerThr	763
Db	1503	ATCCAGAGGGACACCCCTGCTCCAGCCACGTCATCCGGAATCAACACAGCTCCACA	1562
Qy	764	SerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsnTyrThr	783
Db	1563	TCCATCTGGCTCGGTGGAAGAACCCAGATTTCACACAGTCAAGATTGTCACTACACT	1622
Qy	784	ValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyrThrSerSer	803
Db	1623	GTGGCTTACAGCCCTCGGGGCTCAGGAATGCCTCCCTGGTCCACTATTACACAGTTCT	1682
Qy	804	GlyGluAspIleLeuIleGlyLysLysProPheThrLysTyrGluPheAlaValGln	823
Db	1683	GGAGAAGACATCTCATTTGGCGGCTTGAAAGCCATTACCAAAATACAGATTTCAGTGCAG	1742
Qy	824	SerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThrLeuPro	843
Db	1743	TCTCACGGCTGGNATGATGGGCTTTGGCTCTGTGGTGGAGCGCTCCACCTGCCT	1802
Qy	844	AspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSerThrVal	863
Db	1803	GACCGGCTTCCACACCCCATCCGACCTGCGACTGAGCCCTGCATGACCGGTCCACGGTT	1862
Qy	864	ArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyrLeuIleLeu	883
Db	1863	CGGCTGCACCTGTGTGCCCCCAAGAGCCCAAGGGAGATCGTGAGTATCTGATCTCTG	1922
Qy	884	TyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThrGluGlyAsn	903
Db	1923	TACAGCAGCAACCAACACAGCAGCTGAGCACCACTGAGTGGACCTTCTCACCAGCAGGAAAC	1982
Qy	904	IlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLysMetGly	923
Db	1983	ATCTTCAGTGTGAGGTCCATGCTGGAGAGCGACACTCGGTACTTCTTCAAGATGGGG	2042
Qy	924	AlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThrLeuGln	943
Db	2043	GGCGGCAACAGAGTGGGACCTGGGCTTTCTCCCGCTGCAGAGATGATCATCGCTCCAG	2102
Qy	944	GluThrPheSerAspSerLeuAspValHisAlaValThrGlyIleIleValGlyValCys	963
Db	2103	GAGAACTGTACAGCTCGCTGCATGCATCAGTCACGGGCATCATCGTGGTGTCTGC	2162
Qy	964	LeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSerSerHisArg	983
Db	2163	CTGGGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2222
Qy	984	GluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeuTyrThrArg	1003
Db	2223	GAATCCCTCCAGGCTGTCTCTCCAGCCACCCCGGGAGTCCCGGCTGTACTCCAG	2282
Qy	1004	AlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeuValHisPro	1023
Db	2283	GCTCGGCTTGGCCCCCAGCCCGCCAGCTGCCCATGAAATTGGAGTCCCTTGTGCACCC	2342
Qy	1024	ArgProGlnAspTrpSerProProSerAspValGluAspLysAlaGluValHisSer	1043
Db	2343	CATCCCCAGAGCTGTGCTCCCGCCACCTCAGACGTGAGGACAGAGGCTGAAGTGCA	2402
Qy	1044	LeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIleSerTrpAla	1063
Db	2403	CTTATGGGTGGGGTGTCTGAGGCGCGAGTCACTCCAAAGAAAGATCTCTCGGCT	2462
Qy	1064	GlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySer	1083
Db	2463	CAACCAAGCGGCTGAGCTGGCTGGTCTCTGGCAGGCTGTGAGCTGCCCCAG---GCA	2519
Qy	1084	GlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeu	1103
Db	2520	GGCCCCCGCGCTCTGACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2579
Qy	1104	LeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyValLysProSerPro	1123
Db	2580	TTGCTGCAGGCTCTGTGTACGACGCATTAAGGGCAATGGGAGGAGAGTCAACCCCA	2639
Qy	1124	AlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGlyAlaSerLys	1143
Db	2640	GCCTGCAGGAACCCAGGTGGAGCTGAAGTCAATGTCCACTCTGACTTTAGTGCATTAAC	2699
Qy	1144	GlyCysProAspLeuHisLeuGlnAspLeuGluProGluLeuProLeuThrAlaGluThr	1163
Db	2700	GGGAACCTTGACCTCCATCTCAAGACCTGGAGCCCTGAGGACCCCTGCTCCAGAGGCT	2759
Qy	1164	LeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeuGlyArgGlu	1183
Db	2760	CCTGATCTCATCTCGGCTGTGGGATCCAGGGCAGGGGCGCTGCTGGTGACAGGAG	2819
Qy	1184	LeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeuProGluAla	1203
Db	2820	TTGGAGGGTGTGAGCTGGCAGCCCCCGGGCCAGACAGACTTACCTGCTTGTCCAGAGCA	2879
Qy	1204	AlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAlaIleGluAlaPro	1223
Db	2880	GCCAGTGTCTCTGCTCTTACCCGGACCTCCAGCCAGGGGAGGTGCTAGAGGAGACCCCT	2939

QY 1224 GlySerCysGlnProLeuAlaLeuCysProLeuThrValSerProSerLeuProArg 1243
 Db 2940 GGAGATAGCTCCAGCTCCTATCCCTCCCTCTAGGAGCCAGCCAGGCTGCCAGA 2999
 QY 1244 AlaProValSerAla 1249
 Db 3000 TCCCGGCTCTCTCTCT 3017

RESULT 11
 AC110235/c
 LOCUS
 DEFINITION Mus musculus clone RP23-100M12, WORKING DRAFT SEQUENCE, 17 ordered pieces.
 ACCESSION AC110235
 VERSION AC110235.5 GI:20800205
 KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Birren,B., Linton,L., 158055
 Mus musculus, clone RP23-100M12
 Unpublished
 2 (bases 1 to 158055)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArnell,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McSwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlangwa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 158055)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArnell,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McSwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlangwa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (15-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 15, 2002 this sequence version replaced gi:20531900.
 All repeats were identified using RepeatMasker:
 Smit, A.P.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L22383
 Center clone name: 100 M.12
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 153807 bases at least Q40
 Consensus quality: 155134 bases at least Q30
 Consensus quality: 155784 bases at least Q20
 Insert size: 218000; agarose-fp
 Insert size: 156455; sum-of-contigs
 Quality coverage: 7.1 in Q20 bases; agarose-fp
 Quality coverage: 10.0 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and the accession number will be preserved.
 * 6459: contig of 6459 bp in length
 * 6559: gap of 100 bp
 * 6667: contig of 108 bp in length
 * 6767: gap of 100 bp
 * 7863: contig of 1096 bp in length
 * 7963: gap of 100 bp
 * 9426: contig of 1463 bp in length
 * 9526: gap of 100 bp
 * 11011: contig of 1485 bp in length
 * 11012: gap of 100 bp
 * 11112: contig of 1630 bp in length
 * 12841: gap of 100 bp
 * 12842: contig of 1948 bp in length
 * 14790: gap of 100 bp
 * 14889: contig of 2937 bp in length
 * 17826: gap of 100 bp
 * 17927: contig of 3318 bp in length
 * 21344: gap of 100 bp
 * 21345: contig of 6069 bp in length
 * 27414: gap of 100 bp
 * 27514: contig of 8668 bp in length
 * 36382: gap of 100 bp
 * 36481: contig of 8623 bp in length
 * 45105: gap of 100 bp
 * 45205: contig of 13165 bp in length
 * 58369: gap of 100 bp
 * 58470: contig of 27313 bp in length
 * 85783: gap of 100 bp
 * 85882: contig of 26824 bp in length
 * 112706: gap of 100 bp
 * 112807: contig of 41121 bp in length
 * 153928: gap of 100 bp
 * 154027: contig of 4028 bp in length.
 * 154028: Location/Qualifiers
 1. 158055
 /organism="Mus musculus"
 /mol_type="genomic DNA"

FEATURES
 source

```

/db_xref="taxon:10090"
/clone="RP23-100M12"
/clone lib="RPCI-23 Female Mouse BAC"

```

Alignment Scores:	5, 71e-117	Length:	158055
Pred. No.:	Score:	Matches:	913
	3125.50	Conservative:	23.34*
Percent Similarity:	23.29†	Mismatches:	6
Best Local Similarity:	47.33‡	Indels:	3002
Query Match:	2	Gaps:	13
DB:			

US-09-754-997A-2 (1-1252) x AC110235 (1-158055)

Qy	332	AlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAla	351
Db	85019	GCTGCCCCAGCCATCTCGAGGGCCCGAGGGCGCTCTCGCGACGCGGGCCAGCACCGCG	84960
Qy	352	ArgPheValCysArgAlaSerGlyGluProArgProAlaLeuHiTrpLeuHiAspGly	371
Db	84959	CGTTTGTGTGCCGGGGGTTCGGGGAGCCACGGCCCGCGCTGCATCTGGCTGCACGACGGG	84900
Qy	372	IleProLeuArgProHengGlyArgValIleValGlnGlyGlyGlySerLeuValIle	391
Db	84899	ATCCGTTGCGACCCAAATGGCGCGCTCAAGGTGCAGGGCGGTGGCGGCGAGCTTGTGTCA TC	84840
Qy	392	ThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGly	411
Db	84839	ACTCAGATCGGCGCTTCAGGACGCTGCGTACTACAGTGGTACGAGAAAACAGCGCGGGA	84780
Qy	412	ThrAlaCysAlaAlaAlaProLeuAlaValValValArgGluGlyLeuProSerAlaPro	431
Db	84779	ACTGCTGTGCGCGTGGCCCTTGGCGGTAGTGTGTCGCGAGGGGTGTGCCAGCGCCCGC	84720

Qy	524	-----	524
Db	83640	TCCTCAGATCTGGGCTGGGCTACTTCTGACTTGTGTGACCTGGGCAAGTCTCT	83581
Qy	524	-----	524
Db	83580	GATCACTCCCTGACTGTCTTATCTTCCCTGTTGGTGGAGGGAGCTCAGCCAAATGTTT	83521
Qy	525	-----	541
Db	83520	TTCTGTCTCAGTCCCGAGCGCAGCACCAGCTTACCTTGTCCAGCCCAACCCCTCGGA	83461
Qy	541	pIleArgValAlaTrpLeuProSerSerLeuSerAsnGlyGlnValLeuLysTy	561
Db	83460	CATCAGGGTGGCATGGTGGCTCCCTGCTCCCTCCAGCTCAGCAATGGAAGGTCTGAATG	83401
Qy	561	rlyAlleGluTyGlyLeuGlyLys	569
Db	83400	CAAGATAGATACGGTTTGGGGAAGGAAGTGGTGGGCGCGGGAGGTACAGAGATGTC	83341
Qy	569	-----	569
Db	83340	AGTGGGAATGGGGTCTCCCTAGATGAGGAGCGTTGTCACTGGAAGTGGAGCCTAATT	83281
Qy	569	-----	569
Db	83280	GGGGATGGGATGAGAAAGCTCTGGAAGAGCAATGGGAGACTAGATAGCCCAATGTGT	83221
Qy	569	-----	569
Db	83220	AAAGTATGTGTGCAAGCCCAAGGCTGAGTTCAATCTCAGACTTCATGTTAAAAA	83161
Qy	569	-----	569
Db	83160	AAAGCTGGCAGTGTGTAGTGCTCTGTTAATCCAGAGCAGGGAGGTGGAGACAC	83101
Qy	569	-----	569
Db	83100	GGGGACACCTTGAGCTCACTGGGCTAAACCACTGAGTTCCAGGCCAGGAGTCCCTGT	83041
Qy	569	-----	569
Db	83040	GCAGAAGAGAGGTGGCAATACCAAGAAATGACAAACAGAGGTTGTCACTGACCTCCA	82981
Qy	569	-----	569
Db	82980	CATGTATGTGTACACACCTGACATAGTGAACACTCACATCTCATGGACACACAGGA	82921
Qy	569	-----	569
Db	82920	CACAGGGTTTGGGGCAGTTGAGGAAAGGGAAGCAAGAGGTCACTAGCTGGCCAC	82861
Qy	569	-----	569
Db	82860	AGATAAGGAGGTGATCAGTGGGACCTTGGACTGTTCCTTCTGTGAAGGAGCAATTT	82801
Qy	569	-----	569
Db	82800	ACCAGTGGATCCTTGGAAACCTTCTAGGTTTTTGTGTTTGTATTTTAAAGATGTAT	82741
Qy	569	-----	569
Db	82740	TTATTTATTTATATATAGATTCTCTATCTGACGCCCACTGTCATGCCAAGAGGG	82681
Qy	569	-----	569
Db	82680	AATCAGATCCAGTACAGATGGTTGTGAGCCACAGTGTGTCTGGGAATTTGAATCTCAG	82621
Qy	569	-----	569
Db	82620	GACCTCTGGAAGACGTCAGTGTCTCTAACTTTGAGCCATCTCTCCAGCCCCAGAAAC	82561
Qy	569	-----	569

Db	82560	CTTCTGGTTTACACACAGAGAAGATGTGGAGGCTGACAGGATGCATGACATGGG	82501
Qy	569	-----	569
Db	82500	CTGCTCTACTGGGAGATGATGCTGCTCTGGGGAGGTGATGTGTGGAAAGGATG	82441
Qy	569	-----	569
Db	82440	GAGAGTGAGAAAGAAAGTGTGTAGGGTCAAGCTAAGCTAAAGAGAGAGAGCAGAGG	82381
Qy	569	-----	569
Db	82380	CTGATGGAGACAGCAGTCAAGGCTCTCTGACACAGGGACTGCTTAGTGTCTTAGCCAG	82321
Qy	569	-----	569
Db	82320	CCATCAAGCCTGTGTAAAGAGGTGTGTACTGGCTGTGTGTTTTCACAGGGAGGAAAA	82261
Qy	569	-----	569
Db	82260	GTTTCATGCTTGTTCCTCATATGGAACGAGGCCAATGGGCTTCAGCAGCGATAGGTA	82201
Qy	569	-----	569
Db	82200	GAGCCAGGGTCAACCTGGAGGCATGAACCTGCTCCGTCTGAGGGGTCACTGCCCA	82141
Qy	569	-----	569
Db	82140	GAGGACCTCTGGGACCTGAAATGGAAGTCAATGTGATAAATTTTAAAGATTGATGTC	82081
Qy	570	-----	576
Db	82080	CCCAATGAAGCACCCAGCAGCCTATCTTTTGTCCCAAGAGATCAGGTTTCTCCAC	82021
Qy	576	rGluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTy	596
Db	82020	CGAGGTGCTGGAAATGAGACACAACTTACGTTAAACTCACTTCAGCCAAACAAAGTGA	81961
Qy	596	rArgValArgIleSerAlaGlyThrGlyAlaGlyTyGlyValProSerGlnTrpMetGl	616
Db	81960	CCAGTCCCGATTTCACTGAGCTGGCACTGGCGTGGCTATGGAGTCCCTTCTCAGTGGATGA	81901
Qy	616	nHleAsnThrProGlyValHleAsnGlnSerHle	627
Db	81900	GCACAGACACCTGGTGTGCAACCAACAGAGCCATGGTATGGACCTACACGGGCAGACTGG	81841
Qy	627	-----	627
Db	81840	ACAGCGGGACAGGTGCTCTGTGGCCTTAGATATTGCCCTGTTATTGCTGTTATTGCC	81781
Qy	627	-----	627
Db	81780	ACCCCCCCCCCCCCACAAACACACACTTGTATGACCCCGAAGGCTGGGTGTAGAGTGA	81721
Qy	627	-----	627
Db	81720	AGGGCTGGGCTCAAAATATCTCTGTGTTCACTCTGTGACTTTTGAACGCCACTCGCATCTC	81661
Qy	627	-----	627
Db	81660	CAGAGCTTCGGTACTACCAAGGATGATAGCAGGACTAGTCAATTGTAAATTATACACAGAT	81601
Qy	627	-----	627
Db	81600	GGTGGCATAGAACAGAGGCTCCTTGGGTTGAGACCAAGGCTCACTTTCAGAGAGTGTG	81541
Qy	627	-----	627
Db	81540	CCTAGCATGTACAAAGCACTAAATTCATCCCGAGGACCAATAAACCCAGCATACCTCT	81481
Qy	627	-----	627

Db	77101	CCAAATGGTGAGATTGTGGAGTATCTAAATTCTCTACAGCAACACACACCCAGCCGGAAC	77042
Qy	893	ieGlnTrpThrLeuLeuThrThrGlu-----	901
Db	77041	ACCAAGTGAACACTGCTCACCACAGA- GGGAAGGACGCCACCCCTGCTCTGTCCTT	76983
Qy	901	-----	901
Db	76982	TGTTTCCAAGGTCCTCTGGGATTGGAGGCTGAAGCTACTGCTGGGGTCAAAAGTTGGTTT	76923
Qy	902	-----GlyAsnI	904
Db	76922	GCCCAATGGTCTTGGCCCTCTGACCACCTGCTCTCTCCCACTTCCCAACACAGGAACA	76863
Qy	904	lePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLeuMetGlyA	924
Db	76862	TCTTCAGTGCAGAGTCAATGGCCCTAGAGAGTCACTCGGTATTTCTTCAGATGGAG	76803
Qy	924	laArgThrGluValGlyProGlyProPheSerArgLeuGluAspValIleThrLeuGlnG	944
Db	76802	CCCGCACAGAGGTGGGGCTTGGGCCCTTTTCCCGCTTGCAGGATGTGATTACTCTGCAA	76743
Qy	944	luThrPheSer-----	947
Db	76742	AGACATTCTCAGTACCAAGGCAGGAGGAGAGGCCCTTGGACAACCTGGAGGCCCAAG	76683
Qy	947	-----	947
Db	76682	ACCCCTGGATTGGCTCCAGGCCCTGCTTCTGGAAATTAATTTCTTTACTTTGGAA	76623
Qy	947	-----	947
Db	76622	CCCTAAGTGTGGGGATGGAGAAGCAGTGACAGCTTGATAGAGGACTGTCTATAGCTTT	76563
Qy	947	-----	947
Db	76562	GAGGGTCTTTTATTTATTTCTCTGTGTGTGTTTGGCACAGTGATATGCGAAGG	76503
Qy	947	-----	947
Db	76502	TCAGAGGACATCAGAGAGTTGGTTTGTGGGGATCTTGCTTTGCTGGAGGGTTGGGTGA	76443
Qy	947	-----	947
Db	76442	GCTCCTGAGTCTGTGAAGAAGTGTCAATAGGCTGGATAGACACTCACAGGGCTCAAT	76383
Qy	947	-----	947
Db	76382	TAAGATACAAATGAGATCCAGGGCTCTTATTGTAAATTATGAAGATGTGAAGATGA	76323
Qy	947	-----	947
Db	76322	GCCAGTGGGGCATTTAGCCACATAGGACTGAGAGGATGAAGCCCTGAGTCCAAATGCA	76263
Qy	947	-----	947
Db	76262	CACCTGAGTGTGCT	76203
Qy	947	-----	947
Db	76202	GTAATCCTTGTGAGGATGACAGACAGGATATTTCTGGGGCTGAGAGGACTGGGGAG	76143
Qy	947	-----	947
Db	76142	TCAGGTGATGTAGACCTCTGTCTCCACACTACATATACATGCCCTTAGATGAACATATA	76083
Qy	947	-----	947
Db	76082	CAGACATGACACACATGTGTGCACCTAGAGGCCACCTTGTATTATATATATACACCA	76023
Qy	947	-----	947
Db	76022	TGAGAGCCCTTGTGGCAGGAGGAAACAACTTCACTGAAGAGCTGATTCTATAGATCC	75963
Qy	947	-----	947
Db	75962	AGTTTTGTGAGAGTGGCAGGTTCTGTATCAGCAGGGAAGACCGTGTGACTCAGTATGA	75903
Qy	947	-----	947
Db	75902	GGCCCACTAAGATCCCGTGGCTCTCTCTATTCCAATCAGAAGGAGTGAGCTTCCAAACC	75843
Qy	947	-----	947
Db	75842	AGGTGTGAACCAAGGTGTGTTACTGCAACACCTTTGATCTCCGTCTCTGTTATTATGAGCACTG	75783
Qy	948	-----AspSerLeuAspValHisAlaV	955
Db	75782	CTTCTCTTCTGACCTGGAGTCACACTGTTCTCCCCCGCAGACTCTCTTGATGTGCACCGC	75723
Qy	955	alThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysA	975
Db	75722	TCAGGGCATCATCGTGGGTGTCTGCTGGGCCCTTCTCTGCTCTCTGCTGCTGCAATGTGTG	75663
Qy	975	laGlyLeuArgGlnSerSer-----	981
Db	75662	CTGGCCTACGACGAAGCTCCACAGGTGAACCCGGGGAACCACTGAGACCACTGAGTTACC	75603
Qy	981	-----	981
Db	75602	AGGAAGGCTGCAGAGTGGACAAAGGAGCGTGTCTTCTCTGCTGCCCCCGGAAGCCCGAG	75543
Qy	981	-----	981
Db	75542	TATAGATAGGCTATGAGAGCTGAGTGAAGTTCACTGTGCTACATGTGCCCCCATGTGAGGC	75483
Qy	981	-----	981
Db	75482	CCTGGGCGTCATCTCCTCAGTAGACACAAACAAAGTGGATAAATTTATTATTAGGATTTCTGAG	75423
Qy	981	-----	981
Db	75422	ATGAAGGATTGAGGTGTGTGTCTCTCATGAGATGTAGAATAAACTTAGGGAAGCCCGAGGA	75363
Qy	981	-----	981
Db	75362	CTCCCACTCAGTAAACAGCGGGCTGAGTACCTGTGGGCTTCTCATGTGCTTAAATGG	75303
Qy	981	-----	981
Db	75302	AAATGGGCCATCTCGCCATACAGGAAGCGTGCCATCTGAGGGGTGAGCGCGAGTGGACT	75243
Qy	981	-----	981
Db	75242	TGTATTCTCTCTTCTGTTTGTGATTGACACACCCAGGAACCAATCCACTCTTGTGCG	75183
Qy	982	-----HisArgGluAlaLeuProGlyLeuSerSerSerGlyT	994
Db	75182	ACCCATAACTTTTGTCTGTGATCAGGGAAGCCCTGCCGGATTTGTCTCTCCTCAGGCA	75123
Qy	994	hrProGlyAsnProAlaLeuTyrThrArgAlaArgLeuGlyProSerValProAlaA	1014
Db	75122	CCCCAGGAACCCACGCGCTCTACACAGAGCTCGACTTTGGGCCCCCAGTGTCTCTGCTG	75063
Qy	1014	laHisGluLeuGluSerLeuValHisProArgProGlnAspTrpSerProProSerA	1034
Db	75062	CCCATGATTGAGTTCCTCTGTGATCTCTCTCCAGGATTTGTTCCCAACCACTCTGAG	75003
Qy	1034	spValGluAspLysAlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgG	1054
Db	75002	ATGTGAAGACAAAGCTGAAGTACACAGCTTATGGGTGGCAGTGTGTTTTCAGATTGCCGGG	74943
Qy	1054	lyHisSerLysArgLys-----	1059
Db	74942	GCCACTCCAGAGAAGGTGAGACTGTGGAGTCTCTTTCAGGTGAGGGACTCACACGAACCCA	74883

QY 1059 ----- 1059
 Db 74882 TTCCAAATCAACACAGGGGGGACACAGAGACAACTTCTGAGGTGAGATGCGTCGC 74823
 QY 1059 ----- 1059
 Db 74822 GCGCACACACACACACACACACACACACACACACACACAGGCACTGAACC 74763
 QY 1059 ----- 1059
 Db 74762 ATACAGAGAAAGCGTGTCCATACCACCTGGGAGGATCCAGAGCTGTGTTTCGAGAGCA 74703
 QY 1059 ----- 1059
 Db 74702 CCCATACTTCTTCACTTACCTAATATGGAGAAAGAGGCTGTCTCACTTTTGTGTGGC 74643
 QY 1059 ----- 1059
 Db 74642 CAAGTAGGGTGGCCCCAGATAAGGTGAACCCCATAGGCCAGGTGTTCTGGGGGACCA 74583
 QY 1059 ----- 1059
 Db 74582 CCCTGACCAAGTGACAGTCAGTCTTCCCATCTCTTCACTTCTCTGCAATTAGAAAGA 74523
 QY 1059 ----- 1059
 Db 74522 GTCCAGCAAGCTTGGGAGCTTTGGATTTTGCACATCAACTGTTTGTAGGCTTATGGTCC 74463
 QY 1059 ----- 1059
 Db 74462 GCCCTTCCCATTTCCCTACTCACTAGTCTCACTCTCTGCGGGGTGGGGCTTTAGTGTG 74403
 QY 1059 ----- 1059
 Db 74402 GAGTTTAGTGTATCTGGAGATAACAGGGTCCACAGGTCTGGACCTGCATGCCCATAC 74343
 QY 1060 ----- IleSerTrpAlaGlnAlaGlyProAenTrpAlaGly 1072
 Db 74342 TGTCTTCTCTCTCTGAGATCTCTCTGAGTCTCTGAGGAGGAGGACCAAACTGGCGAGGC 74283
 QY 1073 SerTrpAlaGlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuThrArgAla 1092
 Db 74282 TCCTGGGCGAGCTGTGAGCTGCCCCAGGGTAGTGTCCAAAGCGGGCTCTGACCCCGGCT 74223
 QY 1093 LeuLeuProAlaGlyThrGlyGlnThrLeuLeuGlnAlaLeuValTrpAspGly 1112
 Db 74222 CTGCTGCTCCAGCGGAAACGGGCGACACTGCTGCTGCAAGCCCTGGTATATGAGCC 74163
 QY 1113 IleIys ----- 1114
 Db 74162 ATAAAGGTGGTGGTGTCTGTTTGAGAGGGAGGAGGAGGAGGATGGGAGGCGAGGAGG 74103
 QY 1114 ----- 1114
 Db 74102 AAAGTCCCCCTCCATGCACTTCCAGGACAGAAAGTGTATTTGAAAGACCTTGTAGCTTC 74043
 QY 1114 ----- 1114
 Db 74042 CAAGGGGAGTGGTGACAGGAAGCCAGTGTGGAGCTTCTTTACCTGGTCTTAAGTGACT 73983
 QY 1114 ----- 1114
 Db 73982 TACAGTAGACTTACAGAGTATCTGAAGGTGGCCCTTAGCCCTTAGATACGGCCAGCGCCT 73923
 QY 1114 ----- 1114
 Db 73922 CCTTGTCTCTGTGAGGTTCTGGGTCTAGTACTTCTGTGTAGCTGTAGTGAGGCGCAACG 73863
 QY 1114 ----- 1114
 Db 73862 TGTGACATTTCTGGTTTCTCATCTCTGCCAGAGCCTGAGCGTGTCTTCTTGTAGGTGGA 73803
 QY 1114 ----- 1114

Db 73802 ACAGTGTCTTTATCATTAATGGGAGTCAGCCTCTCCCCCGCCCCCAATATATTGGT 73743
 QY 1114 ----- 1114
 Db 73742 CATTTTCTGAAGACCAATTCAGTGTGTCTCACTAGCCTTTTGTCTCACTCTGCCAG 73683
 QY 1115 SerAsnGlyArgIysProSerProAlaCysArgAsnGlnValGluAlaGluValIle 1134
 Db 73682 AGCAACGGGAGAAAGAGCGCTCCCGAGCTGCAGGAATCAGGTGAAGCTGAGGTCAATT 73623
 QY 1135 ValHisSerAspPheGlyAlaSerIysGlyCysProAspLeuHisLeuGlnAspLeuGlu 1154
 Db 73622 GTCCACTCCGACTTCGGTGCATCCAAAGGATGTCTTGACCTCCACCTCCAAAGACCTGGAG 73563
 QY 1155 ProGluGluProLeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSer 1174
 Db 73562 CCAGAGAACCACTGACTGACAGAGACTCTGCCCTTCACGCTCTGGAGCTGTGATCTGTCT 73503
 QY 1175 GlnGlyAlaAspTrpLeuGlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyPro 1194
 Db 73502 CAAGGACGACACTGGCTGGGACGAGCTGGGAGGGTGCCCAACCAACCACTGGGCCA 73443
 QY 1195 GluArgLeuThrCysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGln 1214
 Db 73442 GAGAGGCTCACCTGCTTGCAGAGAGCAGCAGTGCCTCTCTGCTCTCTGCTCAGACCTCCAG 73383
 QY 1215 ProSerThrAlaIleGluGluAlaProGlyIysSerCysGlnProIysAlaLeuCysPro 1234
 Db 73382 CCNAGCACTGCTATAGAGAGGCCCTTGGAAAGCTGCGCAAGCCCAAGGCCCTGTGTCT 73323
 QY 1235 LeuThrValSerProSerLeuProArgAlaProValSerSerAlaGlnValPro 1252
 Db 73322 CTAACAGTCAGCCCAAGCCTTCCAGCGGCCCTGTCTCTCTCTGCTCAGGTGCCCC 73269

RESULT 12

AC106509/c

LOCUS

DEFINITION

AC106509

Rattus norvegicus clone CH230-20218, WORKING DRAFT SEQUENCE, 3

unorderd pieces.

AC106509

AC106509.4

HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 224227)

Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, J., Carter, K., Cavazos, I., Cesar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falle, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C.M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebregiorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Harvey, Y., Havlak, P., Hawes, A., Hamilton, C., Hamilton, C., Hamilton, K.,

Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoques, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

- * NOTE: Estimated insert size may differ from sequence length
- * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
- * NOTE: This is a 'working draft' sequence. It currently
- * consists of 3 contigs. The true order of the pieces
- * is not known and their order in this sequence record is

Db	111454	CAAGCGCATGTGCACACATTCATGCACTAATAAGTCTAAAGAGGTTGTGTAAAGCC	111395
Qy	279	-----	279
Db	111394	CAGGGATGAGGGGCACATCTATCCCAAGCACTCAGAGGCGAGAGCGAGATCTCT	111335
Qy	279	-----	279
Db	111334	GAGTTTCAGGGCAGCGGGACTACTCAGAAACCCCTGCCTCGAAAAATTAGAGGGCAGG	111275
Qy	279	-----	279
Db	111274	GGCGGGAGAGAGACAGAGACAGAGACAGAGACAGAGACAGAGACAGAGACAGAC	111215
Qy	279	-----	279
Db	111214	GACACAGAGAGAGACAGACAGAGACAGAGACAGAGACAGAGACAGAGACAGAGAC	111155
Qy	279	-----	279
Db	111154	GAGAGACAGAGACAGACAGACAGATAGAGACAGTGGGAGCAGGGAAGAGGGGAAG	111095
Qy	279	-----	279
Db	111094	GGACAGAGTTGTATTAATATCTGCCGCCACCAGAGGCCATTTAGCCGTCGTTACC	111035
Qy	279	-----	279
Db	111034	CACCTCCCATCTGCTTTCCACCTAGGGATCAGAGTTAGAAGGAATGGATTCAAGCCAG	110975
Qy	279	-----	279
Db	110974	CGACTGTGTGCAGCTTTGCGGGGTTCCGGAGTGGCGGTGTAGTAGGATCTTTGTAGTTG	110915
Qy	279	-----	279
Db	110914	GAGAGGTTGAGGGGACAGTTCCAGGTCCTCCCTTAACCTACTCTCTCTGTGCACCA	110855
Qy	280	AspGlyLysProIleSerThrAspValIleValLeuGlyArgThrAnLeuLeuIleAla	299
Db	110854	GATGGGAAGCCTATTTCACGGATGATCGTTCTGGCGCGGACCAATCTACTCATCGCC	110795
Qy	300	SerAlaGlnProArgHisSerGlyValTyrValCysArgAlaIleLysProLeuThrArg	319
Db	110794	AGCGCGAGCTCGGCACCTCTGAGTCTATGTCTGCGCGGCCAACAAAGCCCCGACCCGA	110735
Qy	320	AspPheAlaThrAlaIleAlaGluLeuArgValLeu-----	331
Db	110734	GACTTCGCCACTGCCGCTGTGAGCTCCGAGTCTGGTAAAGTCTGGCGCTGTGGGACA	110675
Qy	331	-----	331
Db	110674	TTAAGGACAGGTTTGTGTGTAGAGAGAGAGGGCGATCCTCGAAGACTTAACTTGAGG	110615
Qy	331	-----	331
Db	110614	GGCACGACATGGGAGGAGTCTGAATTACTGATTGATGGTATTGTTGAGGAGCTGGATT	110555
Qy	331	-----	331
Db	110554	CGGTAGCTGGCGCGGGGGGGTGGGGTGGAGGTGGATGGATGGGGCGAAATTGAGAGAGGA	110495
Qy	331	-----	331
Db	110494	ATGAGGAGAGGGGTATTTAAGCCCTGGCTTCTGTCTATGACTAGTGAAGCGTGGGAGAG	110435
Qy	331	-----	331
Db	110434	CAGGAGTGGGGTATTGAGGAGGAGATAGTGAATGGTTTCTGAGAGGGTATGAAGC	110375
Qy	331	-----	331
Db	110374	TCAGGGGCCAGAAATCTGTAGTGTAGCAATTTGGAAAGCTGTAGCAGTAATGTGAGTTCT	110315
Qy	331	-----	331
Qy	331	-----	331
Db	110314	ATGCCAAACCGCACTTACAGCTACAGAGTGAAAGCCCTATCTCAAAACAAACTTATCAAA	110255
Qy	331	-----	331
Db	110254	AAACAAACAAACAAACAAATCAGCTGGAGTGGTGTCTGAGGTGGACAGATCAGAAATTCAA	110195
Qy	331	-----	331
Db	110194	GGCCAACTTTGGGCTACAAAGAGATTTCTCAATACAAACAAACAACTGGGGGCCAGGCTTG	110135
Qy	331	-----	331
Db	110134	CTGAAAGCTGGAGTAACCTGCTGATATAGGAGACAATGGGGAGGGAGTGAAGTTACA	110075
Qy	331	-----	331
Db	110074	GCCTAGAGCTTTAGAGCTGGAAATCTGCTGGGCGCTGGGAGAAAGAGGCCCTGTGSCCGA	110015
Qy	331	-----	331
Db	110014	ATGGGCGAGGGGTGACACAGAGCTGGTCCGGGTGACTAGGAGAGAGGACCTTAGAGTG	109955
Qy	331	-----	331
Db	109954	AGGCTGGGGCGCGGTGTTGGGACACTCGGGGAGGGGTAGGCCCTCTTGTACAGAGTCTCC	109895
Qy	332	-----AlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArg	347
Db	109894	CTCTCCCTCCAGCTCCCGCAGCATCTCGCAGCGCCCGAGCGCTGTGCGGAGCCAGG	109835
Qy	348	AlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeuHisTyr	367
Db	109834	GCACGACCTCGCGCTTCGTTTCCCGCGCTCGGGGAGCCAGACTACGCTGCACCTGG	109775
Qy	368	LeuHisAspGlyIleProLeuArgProAsnGlyArgValIysValGlnGlyGlyGly	387
Db	109774	CTGCACGACGGGCGCCGTTGGGACCAACCGGCGGTCAAGGTGCAGGCGGTGGCGGC	109715
Qy	388	SerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysValAlaGlu	407
Db	109714	AGCTTGTCTCATCTCAGATCGCCTGCGAGACTCTGTGTTACTACCACTGCTGAGCGAG	109655
Qy	408	AnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValGluGlyLeu	427
Db	109654	AACAGCGGGGAACCCCTGTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG	109595
Qy	428	ProSerAlaProThrArgValThrAlaThrProLeuSerSerSerSerValLeuValAla	447
Db	109594	CCCACTGCCCGGCTCGGCTCAAGGCAACCGCGCTGAGCAGCTCCGCTGTGCTAGTGCC	109535
Qy	448	TyrGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyrGlnLys	467
Db	109534	TGGAGCGGCTGAAATTGACAGCGAGCAATCATTTGGCTTCTCTCTTCTACACAAAG	109475
Qy	468	AlaArgGlyVal-----	471
Db	109474	GCAAGGGGTAT-GTCTACTCTTGACCA CGGGACCCAGTTTAGGGCAAGTACGGGATTGAT	109416
Qy	471	-----	471
Db	109415	TAGTAAAGAAATTCCTAATGGTGACCTCTGCAGATGAGGGGAAGGATGAGAGCAATCC	109356
Qy	471	-----	471
Db	109355	TTACAGGGAAGTGTCTTCAGTCTAGGATTAGATTATAGAGTGTACACACACACTCCTT	109296
Qy	471	-----	471
Db	109295	CCCGGAGAGATGCTTTACAGGGTTAAAGGAGCTGTGTCACCCCGAGCTTAACCTGCGAGGA	109236

Qy	471	-----	471	-----	471
Db	109235	GGGGACACTTAGACAGCGCAGGGAACTTCTCTGGCCTATCTCTGGAAATTAACAATGGATA	109176	-----	108099
Qy	471	-----	471	-----	569
Db	109175	GACTCGATGGCAATTTATATTCTCCACCTAGATTGTCTGTGACAAATGTAGCTATACATAG	109116	-----	108036
Qy	471	-----	471	-----	569
Db	109115	ACCAGTAGAATCCCCAGTCTCCAGCTGGTCAGATAGGAGTGGGCTGTGACAGTCCCTTT	109056	-----	107976
Qy	472	-----	482	-----	569
Db	109055	GTCCCCGGTGTACCACTCAGGAGTAGACAAATGTGGAGTACCAGTTTCAGTAAACAAT	108996	-----	107916
Qy	483	AspThrThrGluLeuGlnValArgAspLeuLeuProAsnThrAspTyrGluPheTyrVal	502	-----	569
Db	108995	GACACCAACAGCTGCGAGTTTCGGGACCTGGAAACCAACACACGGAATTCACGTG	108936	-----	107856
Qy	503	ValAlaTyrSerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHisThrLeu	522	-----	107796
Db	108935	GTGGCCTACTCCACAGCTAGGGGCGAGCCGACCTCCAGCCAGCCCTGGTGCATACATG	108876	-----	569
Qy	523	AspAsp	524	-----	569
Db	108875	GACGATGGTAGGACCTCTGAACCTGCAGTGACCTGGGACCCAGAGACACTCCCAAGCAGG	108816	-----	569
Qy	524	-----	524	-----	569
Db	108815	GACAGTGTGGCCCAAGCCCTTCCCTGCTTGTAAATGTCTGTGTGCCCTATAAGCA	108756	-----	107736
Qy	524	-----	524	-----	569
Db	108755	ATGGGCAGAAGTTCTTCTCCAGCTGTAAATAATGGGAAACTGAGGCTCAGAAATGAAC	108696	-----	107616
Qy	524	-----	524	-----	569
Db	108695	GACTTGCAAGCTCAGCAGGACAGCTGGGGCTTGAGCCGAGCTCTCCAGCTCTACAG	108636	-----	107556
Qy	524	-----	524	-----	569
Db	108635	TGCAGCCCTTCTTCTTAATGCTCATGTACATCCAGCCATTTGACTCGAAAGCTTCTGGGC	108576	-----	107496
Qy	524	-----	524	-----	569
Db	108575	CACACTGTTCACTAGCCCTGGGATAGAAATGGATGTCTCTCATCTCTAGCCCTAGGGCTC	108516	-----	107436
Qy	524	-----	524	-----	569
Db	108515	CCCCGAAAGGCTCCCTTCTCTCCTCTGATCCCGGCCCGCCGCTACTTCTGACTTGCTG	108456	-----	107376
Qy	524	-----	524	-----	569
Db	108455	TGTGACCTGGCAAGTCTCTGACCACTCTGACCTTGTCTTCCCTGCTGGTGGAGAGGAGT	108396	-----	107316
Qy	525	-----	536	-----	569
Db	108395	TCATGCCAATGTTCTTTGTCTAGTCCCCAGGCGACACCCAGCTCTCCCTGTCCACG	108336	-----	107256
Qy	537	ProAsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGly	556	-----	569
Db	108335	CCCAACCTTCGGACATCAGGTGGCATGCTGCCCTTCCCTCCAGCTTGAGCAATGGA	108276	-----	107196
Qy	557	GlnValLeuLysTyrIleGlyLeuGlyLys	569	-----	569
Db	108275	CAGGTGGTGAAGTACAAGATAGATGCTGGTTCAGGGAAAGGAAGTGTGATGGGCTCAGGC	108216	-----	107136
Qy	569	-----	569	-----	569
Db	108215	AGTCGACAGATGACTGTGGGAATGGGGGTCTCCCTAGTAGGGGAGTGTGTACGTGGG	108156	-----	107076
Qy	569	-----	569	-----	569

QY	740	-----	740
Db	104857	CCAGCCAGCAGCTGTGTGATGCCACAGCCACCCAGCACACCTCTCTGGCAGGAGTTT	104798
QY	740	-----	740
Db	104797	TCCTGTCTAGACATGCCCTCATTTCTCAAGCTGAGCAGAGCAATAAGGAAAGTTTGA	104738
QY	740	-----	740
Db	104737	AAGAGAACTCCATGTCGATGCGTGATTCCTGGCTCTAATTTTAAATCTTTAGATAGTTT	104678
QY	740	-----	740
Db	104677	TTAAGATCATCCACTTATACCACAGATTTCATCAGATTGTCCCATAGTTTCTTTTAA	104618
QY	740	-----	740
Db	104617	GAITTAATTTATTTATTTATATATACGAGAGTACACTGTACTTGTCTTTTGGACACCC	104558
QY	740	-----	740
Db	104557	AGAAAAGGCATAGGATCCCATACAGATGGTTGTGAGCCACCATGTGTTTGTGGGATT	104498
QY	740	-----	740
Db	104497	TGAATCAGCTCTCAACGACTGAGTCATCTCTCCAGCCCCCATCCCATAAATCTTTAGTAC	104438
QY	740	-----	740
Db	104437	AACACATGGTTTGTCTACCTTGACAATCCATTTCTGACAGTCAGAAGCCAAACCTTGGGT	104378
QY	740	-----	740
Db	104377	TCTAATTTACGCTAGGTGATATTTCTCATCAGCAAGTGTCTTACTAAGTATATACCAGT	104318
QY	740	-----	740
Db	104317	GCCAGCTCTGTTCTGGAAATTAGCAATGAGCAGCATGAATAAGATATCTGTTCTCTGGG	104258
QY	740	-----	740
Db	104257	AAITTTGTTTCATGCTTATCTAATAACAATTTTTCATCACCATAATAGTATCCACAGATT	104198
QY	740	-----	740
Db	104197	TTCTACCCAGCTGTAGTATTAATACTAATCCAACTTCTAAGAACCTCAGAGGGGACC	104138
QY	740	-----	740
Db	104137	AGTGAGATGGTTACAGCAAGTGAAGTGTGGTACAGCCAGATGACCTGGGCTTGTTC	104078
QY	740	-----	740
Db	104077	CCGGCACTCCCAAAATTTCTCTGACCCCCCATGCATATATAGTGGCTCATTTCCCTTC	104018
QY	740	-----	740
Db	104017	CCTCCACACAATAATCATGTGTGATCGAAGACGCGAGTATGAGAAATTTTAAAGAA	103958
QY	740	-----	740
Db	103957	ACAAATGTAAAAAAGGAGAAAGAGAGGAGAAAGAAAGAAAGCAGCTCGAGAT	103898
QY	740	-----	740
Db	103897	GTTGGTGTGTGTCAGATCCATATATAAGATATAGAACAGAGCTACAGAGAAATCCCA	103838
QY	740	-----	740
Db	103837	ACTTCCAAAGCCACATAGTACGTAAGTGGCAGAGAGGAGTTTCAACCAAGGAGCCT	103778
QY	740	-----	740

Db	103777	CGTTCCAGAGTCATGSCATGCTTTCTCTCTGTCTCATATTTTCATCAGTCATGGATGTGTAC	103718
QY	740	-----	740
Db	103717	AGTCACTCTATGTGTGACCTTAGCAAAAAATTCATTTCAACTCTGGAGCCTCAGTTGCA	103658
QY	740	-----	740
Db	103657	GCAGGATTAACATCACCCCACTTACAGGATCAAAAGAGAGGCCGGTGCCAGCTATGGTA	103598
QY	740	-----	740
Db	103597	TCCCACAGCTGTAAACCCCTGTGTCTAGAGGATCAGAGCAAGAGCTAGCCTGGGCTACTAC	103538
QY	741	-----	745
Db	103537	ATAATGAACATCTGCCCATCTCTCTTACCTTGGATTTTCCCTCTAGATCTGCCCTATCC	103478
QY	745	lnArgGlyProLeuProAlaHisAlaHisAlaGluSerAsnSerSerThrSerI	765
Db	103477	AGAGGGGGCCACCGCTGCTCTCGGCCCATGTCCACGGGANTCAACAGCTCCACTCCA	103418
QY	765	leTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsnTyrThrValA	785
Db	103417	TTTGGCTTCGGTGAAGAAGCCAGACTTTTACACAGTCAAGATTGTCAACTACACTGTGC	103358
QY	785	rgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyrThr	801
Db	103357	GCCTTCCGCCCTGGGGGCTCAGGAATGCCCTCCCTGGTCTCACTATATACAGGTATCAGC	103298
QY	801	-----	801
Db	103297	AAGGGAGGGCAACTGGAAATAGTAATACTAGCGGGCTCAGTATGGCTAGGGCACACCCAG	103238
QY	801	-----	801
Db	103237	AATCTACACAGATGTCTCTTTAGAACTGTGTGACCGAGTAGTAACAATGAGGGAGACTGT	103178
QY	801	-----	801
Db	103177	CTGGTCCCACTTAATGACCGAGTGTGTGTCACTTGTGTGTGAGTGGGATGCCATC	103118
QY	801	-----	801
Db	103117	CATTGGCTGGACCTGCTCCCTCCCATCATGAGAGATAGTCTTAAGAACTGCTGTGTCTG	103058
QY	801	-----	801
Db	103057	ATTGAGAAGACGCCCTGTGAAAGCCCAACAGAGAGATAAACTAAGCCAAGGATAGGTAGAT	102998
QY	801	-----	801
Db	102997	ACCTCAGAAGTGACTGGTCCCCCACCCTGGTAAAGAAATACTTAGATTAATTCAGGAA	102938
QY	801	-----	801
Db	102937	GGCATCCCCGAGAGAGAGAGGTCTCTCTTTGGAGATTAGACATGGAAATACCATATGTCCG	102878
QY	801	-----	801
Db	102877	ACCTCTCTGTCTGTTTCCCTGGCTGTGTGTGGGAGGGAGGTAGTCTCCCTGAGTAGCAG	102818
QY	801	-----	801
Db	102817	CACCTTCAGGTTGGGGCTGCTGTGAGCGAGGGGCTTATGTCTGGGGTTCTGCCTGGCT	102758
QY	802	-----	807
Db	102757	TGAGGCTGTCTGAGGGCACTAGTCTATTCCTTTCTCCGTGGAAGCTCTGGAGAGACATC	102698
QY	808	LeuIleGlyGlyLeuLysProPheThrLysTyrGluPheAlaValGlnSerHisGlyVal	827


```

misc_feature      /note="assembly_fragment"
2886..4859
misc_feature      /note="assembly_fragment"
4960..6416
misc_feature      /note="assembly_fragment"
6517..8739
misc_feature      /note="assembly_fragment"
8840..10868
misc_feature      /note="assembly_fragment"
10969..13737
misc_feature      /note="assembly_fragment"
13838..15329
misc_feature      /note="assembly_fragment"
15430..17938
misc_feature      /note="assembly_fragment"
18039..19981
misc_feature      /note="assembly_fragment"
20082..23587
misc_feature      /note="assembly_fragment"
23688..27447
misc_feature      /note="assembly_fragment"
27548..30161
misc_feature      /note="assembly_fragment
clone_end:T7
vector_side:left"
30262..34737
misc_feature      /note="assembly_fragment"
34838..40101
misc_feature      /note="assembly_fragment"
40202..44039
misc_feature      /note="assembly_fragment"
44140..49950
misc_feature      /note="assembly_fragment"
50051..55555
misc_feature      /note="assembly_fragment"
55656..60496
misc_feature      /note="assembly_fragment
clone_end:SP6
vector_side:right"
60597..68544
misc_feature      /note="assembly_fragment"
68645..76940
misc_feature      /note="assembly_fragment"
77041..85480
misc_feature      /note="assembly_fragment"
85581..93905
misc_feature      /note="assembly_fragment"
94006..108063
misc_feature      /note="assembly_fragment"
108164..123496
misc_feature      /note="assembly_fragment"
123597..139965
misc_feature      /note="assembly_fragment"
140066..176282
misc_feature      /note="assembly_fragment"

```

BASE COUNT 44272 a 43276 c 42765 g 43361 t 2608 others
ORIGIN

Alignment Scores:
Pred. No.: 9.29e-94 Length: 176282
Score: 2557.00 Matches: 877
Percent Similarity: 20.85% Conservative: 45
Best Local Similarity: 19.83% Mismatches: 111
Query Match: 38.72% Indels: 3397
DB: 2 Gaps: 18

US-09-754-997A-2 (1-1252) x AC068507 (1-176282)

```

Qy 195 LeuLeuGlnIleLeuAspValGln-----AspSerAspAlaGlySer 208
|||||:|||||
Db 167329 CTCATGGAACCCCTGCAGGTAAAGGCTTAATCCCTTACAGATGGAGACAGCGGGCT 167270
|||||:|||||
Qy 209 TyrArgCysValAlaThrAenSerAlaArgGlnArgPheSerGlnGluAlaSerLeuThr 228
|||||:|||||

```

```

Db 167269 CAGAGAGTCCCAGCAGAGAGAGAAATGAGATTCGACACAGGACGGTGTGTTGCTTT 167210
|||||
Qy 229 ValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIleValAlaAla 248
|||||:|||||
Db 167209 ATCCGGGTGCAGAGTGTCTCGAGTTCGGAGACATGAGGGAAGGCGATCCTACGCTC 167150
|||||:|||||
Qy 249 ProGluAenThrThrValValSerGlyGlnAenValMetGluCysValAlaSerAla 268
|||||:|||||
Db 167149 CCTCCAGC-----GGCTCTCC 167132
|||||
Qy 269 AspProThrProPheValSerTrpValArgGlnAspGlyValProIleSerThrAspVal 288
|||||:|||||
Db 167131 TCTCTTCCCCA-----GACGGGAAGCCCATCTCCACAGATGTC 167093
|||||
Qy 289 IleValLeuGlyArgThrAenLeuLeuIleAlaSerAlaGlnProArgHisSerGlyVal 308
|||||:|||||
Db 167092 ATCGTCTCGGGCGCCACCAACTACTAATTGCCACGCCGACCCCTGGCACCTCCGCGTC 167033
|||||
Qy 309 TyrValCysArgAlaAenLysProLeuThrArgAspPheAlaThrAlaAlaGluLeu 328
|||||:|||||
Db 167032 TATGTCTGCGCGCCACCAAGCCCGCACGCGGACTTCGCCACTGCACGCCGCTGAGCTC 166973
|||||
Qy 329 ArgVal----- 330
|||||
Db 166972 CGTGTCTGGTGAAGCGGGCCGAGGGCGGAGCGGGTTAGCTGGGGCTGGGGTGA 166913
|||||
Qy 330 ----- 330
Db 166912 ATAGACAGGGGAGGTGACCGGTGGGTGTGTGGGGGACATCAGCGCGCTCTCTCGGAG 166853
|||||
Qy 330 ----- 330
Db 166852 GACGAACTCAGGGAGGAGTCCCTAGTAAGGCGGAGGTGGGACGAGGATGGGACTTGGT 166793
|||||
Qy 330 ----- 330
Db 166792 AACTGGAGGACGAGTGCAGGGCGGCTTGAAGTCTGGGATGGAGCGAGGGGTGATTAAGGC 166733
|||||
Qy 330 ----- 330
Db 166732 CTGGACGGGTGATTAGCGGGAGTTGTGACTGGGGTTATTGTGGGGATAGTGGAGTAT 166673
|||||
Qy 330 ----- 330
Db 166672 CTGAGTGGTACTAAGAAACTGGGGGCGAGGTAGATCTGTTAGTGTGGACATATGGG 166613
|||||
Qy 330 ----- 330
Db 166612 AGGGCAGGGAGGAAGGAGATGGGTAATTTAGGGGTGGGTTAGATCCGGGAGGAGCAAT 166553
|||||
Qy 330 ----- 330
Db 166552 ATTTAGGGGAGGCTGAGGAGGCTTGGGGACAGTTTTTGGGAGAGAAAGAGAATGCTGTTG 166493
|||||
Qy 330 ----- 330
Db 166492 TGGGGCTGGGGTAAATGGGGAGGAAGAAAGGGTAGGTCTGCGCTGAGCGGAGTAGTTGG 166433
|||||
Qy 330 ----- 330
Db 166432 AAAGGGTGAATCAGGGAGAGGTTAAAGTGATTGGAGGGGAGCGAGGCGGTGGGG 166373
|||||
Qy 330 ----- 330
Db 166372 GGCAGGCTTGGGGTACTGGGGGGGGGGGACAGAGACTTAATTGGGGATCCCGGCTCAGC 166313
|||||
Qy 331 -----LeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerAr 345
|||||:|||||
Db 166312 GTGCGCCCTCGCGCCCTAGCGCTCCCGCATCCTCAGCGCCCGGCGCTGTGCGG 166253
|||||:|||||
Qy 345 gThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLe 365
|||||:|||||

```



```
Db 166252 GACGGGGGAGACACAGCGCGCTTCGTGTGCGCGCGCTCGGGGGAGCGCGCGCCAGCGCT 166193
Qy 365 uHieTriLeuHisAspGlyIleProLeuArgProAsnGlyArgValIysValGlnGlyI 385
Db 166192 CGCGTGGCTGCACAACAGCGGGCGCGCTCGCGGCCAACAGCGCGCTCAAGGTCAGGGCGG 166133
Qy 385 yGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrGlnCysVa 405
Db 166132 CGGTGGCAGCTGGTCATCACAGATCGGCTCGCAGACCGCGCTACTACCAATGCGT 166073
Qy 405 lAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValArgG1 425
Db 166072 GGCTGAGAACAGCGGGGAGATGGCGTGCCTGCCGCTCGCTGGCGGTGGTGGCGGA 166013
Qy 425 uGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValle 445
Db 166012 GGGGCTGCCAGCGGCCCAAGCGGGTCACTGCTACGCCCACTAGCAGCTCCGCTGTGT 165953
Qy 445 uValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTy 465
Db 165952 GTTGGCCTGGGAGCGGCCGAGATGACAGCGAGCAGATCATCGGCTTCTCTCCACTA 165893
Qy 465 rGlnIysAlaArgGlyVal 471
Db 165892 CCAGAGGCAAGGGGTAC-GTCCGGCTGGGGCGGATGCAACAGCGCAGGCCCGGGCGGG 165834
Qy 471 471
Db 165833 GCGCAGCCAGAAGGGGTTCAAGTCGGATGATGTAAGGGTTGCTGTATGGTACCTCTGG 165774
Qy 471 471
Db 165773 GCAAGGATGGACTGGGGTAGGAGCCCATAGGAGAGACCCGCATAGTATGGTAGGTATGG 165714
Qy 471 471
Db 165713 ATTAGGCAACTGGCGCTGGAAGGGCCAATGCTTCCCTCAGTCATGTTTGTGAGAGAGGA 165654
Qy 471 471
Db 165653 GGGAGAGGACGTTCTCTGGAGCCACCTCCATCCCAAGTTAACTAATCTGGAGAGGGGGTCA 165594
Qy 471 471
Db 165593 TTTCCAGAAAGCACTCCCTTAACCAAGTGGAGGAATTAATACATGAGACAGGAAGCTGAGG 165534
Qy 471 471
Db 165533 GAGTTTATACACTCAGCTCAGATTATCTGGGATGATTTAGCTATAAATAGACACAGGA 165474
Qy 471 471
Db 165473 CAATATGAAGGATTCCTATACAGACTCCAGCTAGCCAGCAGCGGGGGTGGCACTATGAC 165414
Qy 472 472
Db 165413 AGTCCCTTGTCCCGGTGTACACAGGCAATGGCAATACCAAGTTTGCAGTGAA 165354
Qy 481 nAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTy 501
Db 165353 CAACGACACCCAGAACTACAGGTTCCGGACCTGGAACTCCCAACACACAGATTATGAGTTCTA 165294
Qy 501 rValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerProAlaLeuAlaHisTh 521
Db 165293 CTTGTGGCCTACTCCCAAGCTGGGAGCCAGCGCACCTCCACCCCGCAGCACTGGTGACAC 165234
Qy 521 rLeuAspAspVal 525
Db 165233 ACTGGATGATGG-TAGGGCTCTGAACCTCGCAGTGGCAGCTTGGGGTAAAGAGAAATTC 165175
Qy 525 525
Db 165174 CTGGGAGGGGAGCTCACCCCTCTCTTACTAATGTAAGGCAATTTTACTCTCATATC 165115
```

Qy	570	-----	570
Db	164034	GSTGGCCCTCAGGGCTCTGGAAGCCAGCGGTTCCTTGCTGGCCTTCAATGGTCAGCCTCTTG	163975
Qy	570	-----	570
Db	163974	TGTCTGCATCCCCTGTGTGTGTGTGTGTGGCAGGGGATGTCATTGACAGCTCCTGGGGAA	163915
Qy	570	-----	570
Db	163914	GTGGGGAAGCCCTGCCATCATTTCTCCCTCAGCTGAGCCAGGCTGATGGCTGCATCAG	163855
Qy	570	-----	570
Db	163854	CCAGAGTTGGAGGCTCAAGCGTCAAACCTCGAGCCATAGAATCTGCCAGGCACTGAGGG	163795
Qy	570	-----	570
Db	163794	GTCAAGTCTAGAACGCTCAGGCTCCACGGGCAAGGAGAACCGGATGTAGGCGGA	163735
Qy	570	-----	570
Db	163734	AGAGCCATGTAGGGGGTTTTAGGATCCTGTCTCAGTGGCTACCAGCAGCCCCACCTT	163675
Qy	571	-----AspGlnValPheSerThrGluValProGlyAseGluThrGlnLeuTh	586
Db	163674	TTCTGTCCTTCAGATCAGATTTCTACTAGGTGCGAGGAATGAGACACAGCTTAT	163615
Qy	586	rLeuAenSerLeuGlnProAenIysValTrArgValAlaGllleSerAlaGlyThrDlval	606
Db	163614	GCTGAATCTGCTCAGCAACAAGGTGTATCAGTACGGAATTCGGCTGGTACAGCAGC	163555
Qy	606	aGlyTr-GlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAseGlnSe	626
Db	163554	CSSCTTCGGGGCCCCCTCCAGTGGATGCATCACAGACGCCAGTAGTATGCACACACAG	163495
Qy	626	rHis-----	627
Db	163494	CCATGGTACGGGGTTAGAGGAGAGACAGGACGTGTGTGTGACTGTGTGGGTGTGGT	163435
Qy	627	-----	627
Db	163434	GTGTCTTTGGGACCCAGCGCCCAACCCCTGGCCCCCTCCATCAGAGACCTTCCATG	163375
Qy	627	-----	627
Db	163374	GCTCCCTCCCTTGATGGGTGAGTGAGGCTTAGTCCAGAGGTGAAGGCTGGGATCTGG	163315
Qy	627	-----	627
Db	163314	TGTCCCTTCGATTGTCTGTGACCTTGGACAAACCAATTCGTAGGGCTTCATTACTGCC	163255
Qy	627	-----	627
Db	163254	AGGAATGGCATGAGGTGAATAATGTAATTGTAACTAAAGTGCTTAGTACATAAAAAC	163195
Qy	627	-----	627
Db	163194	TTGAATGTAGCTCTCAGAGGATGATGGTGATGATGATGTTGTTTGTGAGAGGCCNA	163135
Qy	627	-----	627
Db	163134	TAGAAGAGTGATTAAATGCTGACCAACCGTCTACATGCTCTCTATGCGACACTAGGTA	163075
Qy	627	-----	627
Db	163074	AATTACTTAATCTCTGAGCTTCAGTTTCTTATCTGTAAAGCGAAGTTGTAGTATTTT	163015
Qy	627	-----	627
Db	163014	CACTATTTTAGAGGTTGCAAGGATMAAGTTTAAATACTACAAAGTGATAGACAG	162955
Qy	627	-----	627

Db	161876	CCTAGGATCAGGGTCTCTTTTAGCTCCAGAGAGCCTCTGTCAATGAGAGAGCTTGT	161817
Qy	741	-----	741
Db	161816	TGTTCACTGCTTAGGCTGCCACCTTATTTGCAATTACCCATGAATAGTCTGATTTCCA	161757
Qy	741	-----	741
Db	161756	AACTTTGACAGAGGGAACCTTATGTTGATCAGTAACCTCCGACCCCAAGTTTGACTG	161697
Qy	741	-----	741
Db	161696	TAAACAAAGTTTTTTTAAACAAACATTTGTCTACATCACACTATCTCTCAAG	161637
Qy	741	-----	741
Db	161636	TCTAAGAAGAGTGAGTATCCTCTGTGACCTAACGGGAACCTGAGGCTCAGAGAGGA	161577
Qy	741	-----	741
Db	161576	GGAACCTTGTCCAAGGTCACACAGAGCCTAGGCTACAACTGGGACTCTGTGTTTCCAGGCA	161517
Qy	741	-----	741
Db	161516	AAACTCTTCCACTGCATCACTGCTGTGGCCCTGCAATGGACACCTCAGTGGCCCA	161457
Qy	741	-----	741
Db	161456	TCATGATACCTCCCTCACCCCTAACTTCCAGCCACATGGTTTAGCAGAGGAGTAATCAA	161397
Qy	741	-----	741
Db	161396	CCAGAAGCCATGAAGCTTGGGTTCTTTTCTTTTGTGAGACAGAGTCTCACTCTGTG	161337
Qy	741	-----	741
Db	161336	CCAGGCTGGAGTGCAATGGTATGATCTCAGCTCACTGCAACCTCTGCCCTCTGGGTCA	161277
Qy	741	-----	741
Db	161276	AGCAATCTTGTGCTCGGCTCTTGAGCAGCTGGGATTAACAGGTGTGCACCATACCTGG	161217
Qy	741	-----	741
Db	161216	CTAATTTTGTATTTTAGTAGAGCGGGTTTACCCTGTTGCCAGGCTGGCTTGAA	161157
Qy	741	-----	741
Db	161156	CTCCGACCTCAGGTGATCACCCTCTTGGCCCTCCCAACGTGCTGGATTACAGGTGTG	161097
Qy	741	-----	741
Db	161096	AGCCACCAACCAGCTGATCTAATTCAGTTTGTTCATTTCTTCATTCAGCAATAT	161037
Qy	741	-----	741
Db	161036	TTCTGCTAGTGTCTACTAAGTGCCAGCTGCTTTGGGNAATTTAGTGGTGGACACAAAT	160977
Qy	741	-----	741
Db	160976	AAGGTTCTTGTCTCTTGGGTTTATTTATGTTTCTAATAATAACACTTATAAATATAG	160917
Qy	741	-----	741
Db	160916	TGTTTATTAATTTCAATTTCTTATCTATTTTAAAAATATAAATAATTTTAATCTCTATA	160857
Qy	741	-----	741
Db	160856	ACAACCCCATGAGTAGGTACTATCATTAAGATAGTAATGGAGACACAGAGAAGGTAG	160797
Qy	741	-----	741
Db	160796	TAATTTCTCCAAGGTCACACAGCTAGTAAGTGTACAGAGCTGGAATTTGAACCAAGCAGC	160737
Qy	741	-----	741
Qy	741	-----	741
Db	160736	CTGGTTCCAGAGTCTGGTTCCATGGTTATTCATTCGTGTGATTTTGTCTCTCTCTTTA	160677
Qy	741	-----	741
Db	160676	TTACGTTTTCACAGACTAGTGTTCACAGTCAAAATCAAATTTCTTGGAGCCTCAATTG	160617
Qy	741	-----	741
Db	160616	CCTTATCTGTAAGTAGGAATGATGCCACTCAACTCACAGGATCAAAAGGATGGGTCAAA	160557
Qy	741	-----	741
Db	160556	GAAACAGCCTTGGGAACCTGTAACCACTATACATATTAAGGATGATATCTCCCTCC	160497
Qy	742	-----LeuProIleGlnArgGlyProProIleuPr	751
Db	160496	TTCTCTCACACTCTAGTGTTCACCTTTTCAGACATGCTCTATCCAGAGGGGACCCCTGCC	160437
Qy	751	oProAlaHisValHisAlaGluSerAsnSerThrSerIleTrpLeuArgTrpIleVal	771
Db	160436	TCAGGCCACGTCCTATGCGAATCAACAGCTCCACATCACTGCTTCGGTGGAAAAA	160377
Qy	771	sProAspPheThrValIleValAsnTyrThrValArgPheGlyProTrpGlyLe	791
Db	160376	GCCAGATTTCCACACAGTCAAGATTGTCAACTACACTGTGTGGCTTCAGCCCCCTGGGGCT	160317
Qy	791	uArgAsnAlaSerLeuValThrTyrThrSer	802
Db	160316	CAGGAATGCTCTCTGTGCTCACCTATTACACCAG-GTGGGACTAATTCCTTTACTAATGT	160258
Qy	802	-----	802
Db	160257	AAAGAGGGAGGAAGAAATAGTGGGGGGTGGTGGGAGGACAGCAGGTGGTGCATCAGG	160198
Qy	802	-----	802
Db	160197	GAAGGCCCTCTGGATTTCAGGAACATGTCTATGGCTCACCTGTATCGCTGAGGCCACACAG	160138
Qy	802	-----	802
Db	160137	AGGGGCTGCACAGGCTCCACGCTCTTGGCTGTGACAGATGTAGCAAGAGGCCATCC	160078
Qy	802	-----	802
Db	160077	AATCCCGTCCCTCAACACAGAGCCTTGTTCATAACTCCGGAGATGTGACACAGGACA	160018
Qy	802	-----	802
Db	160017	GACCTGTTTCTCCACCTCAGAGGGCTCCAGTCTGTATGGTGAAGAGAACCCACCCCTG	159958
Qy	802	-----	802
Db	159957	TCCTCAGAGCTGCTCCCTCTACCTCTGTAGATGGGGAGACAGAGCCTGAGGAGCCTGAGG	159898
Qy	802	-----	802
Db	159897	TTTGGGGAGGAAGTGCATATAAAGCAACCGATATATACAATGCTGCAATGAAGGCCAAAA	159838
Qy	802	-----	802
Db	159837	TTAAGCCAAAGAAAGATGCATCGGAGATGATTTCCATAGGTGGAGAAATACTTGGATTAA	159778
Qy	802	-----	802
Db	159777	TCCAGGAAGTTTCTCGGAGGAAGGGTCTGGTTTCGAGATTGCACTTGGAGCAGCATAC	159718
Qy	802	-----	802
Db	159717	ATCCAGACTTCTCTGTAGGTCCCTGCGGTGGTGTAGTCCGCTCCAGGCCAGGAGCAGCT	159658

Qy	802	-----802	
Db	159657	TCCCTTCCACGGGTAGAGCCTGGGGCTGGGGCTGCACTGCAGGGGAGTAGATTGGAGGAG	159598
Qy	803	-----Se 803	
Db	159597	TCGTCCCTGCTGGGAGAGTGTGCTGAGAGCACTAGTCACGTTTCTTTCAGCAGTTTC	159538
Qy	803	xGlyGluAspIleuIleGlyGlyLeuLysProPheThrLysTyGluPheAlaValG1	823
Db	159537	TGGAGAAGACATCTCATTCGCGGCTTGAAGCCATTCCACCAATACAGATTTCGAGTGCA	159478
Qy	823	nSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThrLeuPr	843
Db	159477	GTCTCAGGGGTGGACATGATGGGCTTTTCGGCTCTGTGTGGAGCGCTCCACCCCTGCC	159418
Qy	843	oAspAlq-----845	
Db	159417	TGACCGTGTAGTGGCTCCAGCCCTTGCCACCAACCCATCTTTTGGCCCACTCCTTCTTGG	159358
Qy	845	-----845	
Db	159357	CCCCAGCAGCAAGAAACATACTTCAGAACCCCAAGGTTTGGGCAGAGCTGGGGGCAGAAG	159298
Qy	845	-----845	
Db	159297	AGTTTCTGGAGGGGAGGAATGTTAAGTCCCTCCAACACACATACTGTTTCTAGAAAT	159238
Qy	845	-----845	
Db	159237	CCCCAACTAGAAACCAAGACCTCATCTTTGTGTGGCTTGTTCATGAAAGGAGGAAGAA	159178
Qy	845	-----845	
Db	159177	GGGAGGTCAATCATCATGCTCTGTCCCCCAGCTCTGACATCTCTCCCCATCTCCTCTCC	159118
Qy	846	-----ProSerThrProProSerAspLeuArgLeuSerProLeuThrProSe	861
Db	159117	CGCACCCCGAGGGCCTCC-ACACCCCATCCGACCTGCACCTGAGCCCTGCACACCGTC	159059
Qy	861	rThrValArgLeuHisTrpCy8ProThrGluProAsnGlyGluIleValGluTyrlle	881
Db	159058	CACGGTTGCGCTGCACCTGTGTGCCCCCAGAGCCCAACGGGAGATCGTGAGTATCT	158999
Qy	881	uIleLeuTySerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThrG1	901
Db	158998	GATCCTGTACAGCAGCAACACACGACGCTGAGCACCACTGACCTTCTCACCAACGCA	158939
Qy	901	u-----901	
Db	158938	GGGTGAGGGCTCTGTGCCCGAGGGCATCTGCTCAAGAGCTTCCAGCCCTCTCCGCTTCC	158880
Qy	901	-----901	
Db	158879	AGCAGTGAGAACAGTCTGGGGTTTGTTCAGGGGCTACTGGGAGAGCAAGCAGAGCT	158820
Qy	901	-----901	
Db	158819	GCTGCCTAAGTGGAGACAGACTCTCCAGTCCCTCCCTGCCACCTCTGCATCCACCC	158760
Qy	902	-----GlyAsnIlePheSerAl	907
Db	158759	ACCCATCCACCATCCCTGCTCTCCCTGCTCCCTCACCCCGAGGAACATCTTCAGTGC	158700
Qy	907	aGluValHisGlyLeuGluSerAspThrArgTyPhePheLysMetGlyAlaArgThrG1	927
Db	158699	TGAGGTCATGGGCTGGAGAGCAGACCTCGTACTTCTTCAAGATGGGGCGCGCACAGA	158640
Qy	927	uValGlyProGlyProPheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSe	947
Db	158639	GGTGGGACCTGGGGCTTTCTCCCGCTTGCAGGATGTGATCAGCGTCCAGGAAGCTGTC	158580
Qy	947	r-----947	

Qy	209	TyArgCyValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSerLeuThr	228
Db	88616	CAGAGAGTCCACGACGAGAGGAGAAATGAGATTCCGACGACGCGTGTGTTGCTTT	88675
Qy	229	ValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValleAlaAla	248
Db	88676	ATCCGGGTGCAGAGTGTCTCAGAGTTCGAGACTGAAGGGAAGGCGAGTCCCTAGGCTC	88735
Qy	249	ProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCyValAlaSerAla	268
Db	88736	CCCTCCAGC-----GCGTCTCCC	88753
Qy	269	AspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThrAspVal	288
Db	88754	TCTCCTTCCCA-----GACGGGAAGCCCATCTCCACAGATGTC	88792
Qy	289	IleValLeuGlyArgThrAsnLeuIleAlaSerAlaGlnProArgHisSerGlyVal	308
Db	88793	ATCGTCTGGCGCGCACCAACTACTAATTGCAACGCGCAGCCCTCGCATCTCCGGGCTC	88852
Qy	309	TyrValCyArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAlaGluLeu	328
Db	88853	TATGTCTGCGCGCCACCAAGCCCGCACGCGGACTTCGCCACTGCAGCCGCTGAGCTC	88912
Qy	329	ArgVal-----	330
Db	88913	CGTGTCTGGGTGAGCGCGGCCAGAGCGGAGCGGGTTAGCTGGGCTGGGGTGA	88972
Qy	330	-----	330
Db	88973	ATAGACAGGGGAGGTGCACGGGTGGGTGCTGGGGGCACTCAGGCGCTCTCCTGGAG	89032
Qy	330	-----	330
Db	89033	GACGAACTCAGGAGGAGTCCCTAGTAAAGCGGAGGTGGGACGAGGATGGGACTTGGT	89092
Qy	330	-----	330
Db	89093	AACTGGAGGACGAGGTGCGGGCGTTGAGTCTGGATGGAGCGAGAGGGGTGATTAAAGC	89152
Qy	330	-----	330
Db	89153	CTGAGCGGGTGATTAGCGGGAGTTGTCACTGGGTATTCTGGGGATAGTGAGTGAT	89212
Qy	330	-----	330
Db	89213	CTGAGTGGTACTAAGAAACTCGGGGGCAGGGGTAGATCTGGTAGGTGGGACATATGG	89272
Qy	330	-----	330
Db	89273	AGGACGGGAGGAGGAGATGGGTAAATTAGGGGTGGGGTTAGATCCGGGAGGAGCAAT	89332
Qy	330	-----	330
Db	89333	ATTAGGGGAGCTGAGGAGCCCTGGGGACAGTTTTCGGAGAGAAGAGATGGTCTGTG	89392
Qy	330	-----	330
Db	89393	TGGGGCTGGGGTAAATGGGGGAGAAAGGTTAGGTCTGCGCTGAGCCGAGTAGTTGG	89452
Qy	330	-----	330
Db	89453	AAAGGTGAATCAGGGAGAGGTTAAAGTGAATTGGAAGGGAGCGAGGGAGGGCGTGGG	89512
Qy	330	-----	330
Db	89513	GGCAGGTTGGGTTACTGGGGGGGGCAGAGACCTTAGTTGGGGATCCCGCGCTCAGC	89572
Qy	331	-----LeuAlaProLalleSerGlnAlaProGluAlaLeuSerAr	345
Db	89573	GTGCGCCCTCGCGCCCTACGGCTCCCGCATCACTCAGGCGCGCCGAGCGCTGTGCG	89632

Qy	345	gThrArgAlaSerThrAlaArgPheValICysArgAlaSerGlyGluProArgProAlaLe	365
Db	89633	GAGCGCGCGGAGCACACAGCCGCTGCTGTGCGCGCGCTCGGCGGAGCCGCGCGCAGCGCT	89692
Qy	365	uHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValIysValGlnGlyG1	385
Db	89693	GCCTCGCTGCACACACGGCGCGCCGCTGCGGCCAACCGCGCGCTCAAGGTCCAGGCGCG	89752
Qy	385	YGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysVa	405
Db	89753	CGGTGGCAGCCTGGTTCATCACACAGATCGGCTGCGAGGACCGCGCTACTACCAAGTGGT	89812
Qy	405	lAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgG1	425
Db	89813	GGCTCGAAACAGCGCGCGGAATGCGCTGCGCTGCCCGCTGCTGGCGGTGTGTGCGCGA	89872
Qy	425	uGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLe	445
Db	89873	GGGCGCTGCCAGCGCGCCCCACCGCGGTCACTGCTACGCACCTGACAGCACTCCGCTGTGT	89932
Qy	445	uValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisT	465
Db	89933	GGTGGCTGGGAGCGCGCCAGATGACACAGCAGCAGATCATCGGCTTCTCTCTCCACTA	89992
Qy	465	rgLnLysAlaArgGlyVal-----	471
Db	89993	CCAGAGGCACCGGGTAC-GTCGGCTGCGCGCGGATGCACAGCGCAGGCCCGCGCGG	90051
Qy	471	-----	471
Db	90052	GCGAGCCAGAGGGGTTCAGTTCGGATGACTGAAGGGTTGCTGATGGGTACTCTGCG	90111
Qy	471	-----	471
Db	90112	GCAAGNATGCATGGGGTAGGAGCCCATAGGAGAGACCCGCATAGTATGGTAGTATGG	90171
Qy	471	-----	471
Db	90172	ATTAGGCAACTGGCGCTGGAAAGGGCCACTGCTTCCCTCAGTCATGTTTGTGAGAGAGGA	90231
Qy	471	-----	471
Db	90232	GGGAGAGGACGTTCTCGAGACCACTCCCATCCAGTTAACTAACTGGAGAGGGGGTCA	90291
Qy	471	-----	471
Db	90292	TTTCCAGAAAGCACTCCCTAAACAGTGGAGGAATTAATACACTGAGACAGGAGCTGAGG	90351
Qy	471	-----	471
Db	90352	GAGTTTATACATCCAGCTCAGATTATCTGGGATGATTTAGCTATATAAATAGACACAGGA	90411
Qy	471	-----	471
Db	90412	CAATATGAAGGATTCCTATACAGACTCCAGCTAGCCAGCACCGGGGGGTGGCACTATGAC	90471
Qy	472	-----AspAsnValGluTyrGlnPheAlaValAs	481
Db	90472	AGTCCCTTGTCCCCGTGTCAACAGCGCATGGACATGTGGAAATACCAAGTTTGCAGTGA	90531
Qy	481	nAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheT	501
Db	90532	CAACGACACACAGAACTACAGGTTCGGGACCTGGAAACCAACACAGATTATGAGTTCTA	90591
Qy	501	rValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisTh	521
Db	90592	CGTGGTGGCTTACTCCAGCTGGGAGCCAGCGCACTCCACCCAGCACTGGTCACAC	90651
Qy	521	rLeuAspAspVal-----	525
Db	90652	ACTGGATGATGG-TAGGGCCCTCTGAACCTGCGAGTGGGCGAGCTTGGGGTAAAGAGAATTC	90710
Qy	525	-----	525

D	b	90711	CTGGGAGGGAGGCTCACCCCTCTCTTTTACTAATGTAAAGGCATTTTTACCTGCTCATATC	90770
Qy		525	-----	525
D	b	90771	ATCTTTTGTCGCCCGTAAGAGAGTAGGGGGGTTCTTCTCCTCAGTTCACAGTTGGGGA	90830
Qy		525	-----	525
D	b	90831	ACTGAGCCCGAAAAATGAAATGATTTCGCCAGGTCAGTAGTTCAGCTGGGACTTAAGC	90890
Qy		525	-----	525
D	b	90891	CCAGGTCGTGTGATGATTCCTCTGTGTCATGCACTTCTCCCTGACACTATGTGCCATCCAG	90950
Qy		525	-----	525
D	b	90951	CTGGGTAGGGCCAGTGACATTAAACCATTAGACGTGGAAGATTGTCTCCAAGACTTT	91010
Qy		525	-----	525
D	b	91011	TTCAGCTTCAGGGTGGGAAGAAGTGCCTCATCTTCAGCTTGGGCTCCCCCTGAGAA	91070
Qy		525	-----	525
D	b	91071	GCTCCCTTCTCCCTCAGATCCTGTGCTCTGGGCTACTTCTGACTTGTGTGTGACCCAG	91130
Qy		525	-----	525
D	b	91131	GCAAGTCTCTGACCGCTCCGAGCCTCACCTTCCAGTCGGTGGAGCAGAGCTCATGCTG	91190
Qy		526	-----ProSerAlaIaLapProGlnLeuThrLeuSerSerProAsnPr	539
D	b	91191	ATGTTCTTCTGCTCAGTCCCAGTCAGACCCAGCTCTCCCTGTCCAGCCCAACCC	91250
Qy		539	oSerAspIleArgValalaTrpLeuProLeuProSerSerLeuSerAsnGlyValle	559
D	b	91251	TTCCGACATCAGGGTGGCTGGCTGCTCCCTGCCCCCAGCCTGAGCAATGGCAGGTGGT	91310
Qy		559	uLysTyrlsIleGlutryGlyLeuGlylysGlu-----	570
D	b	91311	GAAGTACAAGATAGAATAAGTTGGGAAAGGAAGGTGAGTGGGGTGGGTGGCTGCAG	91370
Qy		570	-----	570
D	b	91371	GCATGTATGGGAATGATCAACGTGAAGAGCTGTACAGGGCTGCTGGAGACCTGAGCAGGT	91430
Qy		570	-----	570
D	b	91431	ACGGAGTCGGAGGAAGGAAGCAGAGTAGGGTTGGGGAAGGTGAACCTGAAAACAG	91490
Qy		570	-----	570
D	b	91491	GGGTCTGGCAGGAATGAATAAGGAAGACAGAGGGCCTAGCTTAAGGCAGAGGGAGGCC	91550
Qy		570	-----	570
D	b	91551	TGGGCTCTACCCCAACCCCTTACAGCTGCAGTGTGTACTTCTACTGAGTCCCTGCCTC	91610
Qy		570	-----	570
D	b	91611	AGTTTCTCTTATCTGGGAATAGGGGGCTGAGGTGGGAGGATCACTTGACCCAGGAATTC	91670
Qy		570	-----	570
D	b	91671	AAGGCTGACGGAACTGTGATCGTGCCATTGTACTCCCATCTGGGTAAACAGAGCAAAACT	91730
Qy		570	-----	570
D	b	91731	CTGCTCAGAAAAAATAAAGGNTAAAAGTTAGGAGAGGAAGGTAGCTGAGGAGCA	91790
Qy		570	-----	570

Db	91791	GGGGAAGGTGAGGAAGGAGGGAAGGTGACGAGGAAGGGAAGGTGAGGAAGTGCCCAT	91851
QY	570	-----	570
Db	91851	GGTGGCCTTCAGGGCTCTGGAAGCCAGGGCTTCTTTGCTGGCCCTTCATGTCAGCCTCTTG	91910
QY	570	-----	570
Db	91911	TGTCTGCATCCCCTGTGTGTGTGTGTGGCAGGGGATGGTCATTGACAGCTCTCGGGAA	91970
QY	570	-----	570
Db	91971	GTGGGGAAGCCTGCCATCAATCTCCCTGCAGCTGAGCCAGGCTGATGGCCCTGCATCAG	92030
QY	570	-----	570
Db	92031	CCAGAGTTGGAGGCTCAAGCGTCAACCCCTGGAGCCATAGAATCTGCCAGGCACTGAGGG	92090
QY	570	-----	570
Db	92091	GTCA GTGCTGAGAACGCTCAGGCTCCACGGGCA CGGAGAACGGGATGTGAGGCGGA	92150
QY	570	-----	570
Db	92151	AGAGCCATGTAGGGGGTTTTAGATCCTGTTCTCAGTGCGCTACCCAGAGCCCCACTT	92210
QY	571	-----AspGlnValPheSerThrGluValProGlyAasnGluThrGlnLeuTh :::	586
Db	92211	TTCTGTCCCTGCAGATCAGATTTTCTCTACTAGGTGCGAGGAATGAGACACAGCTTAT	92270
QY	586	rLeuAasrLeuGlnPröhenLyvValTyArgValArGlIeserAlaGlyThrGlyAl :::	606
Db	92271	GCTGAATCTCGCTTCAGCAAACAAGGTGTATCGAGTACGGAATTCGGCTGTGTACAGCAC	92330
QY	606	aGlyTyGlyValProSerClnTrpMetGlnHlaargThrProGlyValHisAasnGlnse :::	626
Db	92331	CGGCTTCGGGGCCCCCTCCCGATGGATGCATCACAGAGCCCCCATGTATGCACAACACAG	92390
QY	626	rHis- 	627
Db	92391	CCATGGTA CCGGGTTAGAGGAGAGACAGGACAGTGTGTGTGTGACTGTGTGGGTGTGCGT	92450
QY	627	-----	627
Db	92451	GTGTCTTTGGGACCAGCGCCCCCACCCCTGGCCCCCGCCCTCCATCAGAGACCTTCCATG	92510
QY	627	-----	627
Db	92511	GTCCTCCCTTGATGGGTGAGTGAGGCTTAGGTGCAGAGGTGAAGGCTGGGATCTGG	92570
QY	627	-----	627
Db	92571	TGTCCTTCGATTTGCTGTGTGACCTTGGACAACCAATTCGAGGGCTTCATTACTGCCC	92630
QY	627	-----	627
Db	92631	AGGAATGCCATGAGTTGAATAATGTAATTGTAACATAAGTGCTTAGTACATAAAAACCTCA	92690
QY	627	-----	627
Db	92691	TTGAATTTGTAGTGCTTCGAGGATGATGTTGTATGATGATTTGTTTTGTGAGAGGCCAA	92750
QY	627	-----	627
Db	92751	TAGAAGAGTGATTAATGTGACCAACCGGTCTACATGCTCTCTATGGCACACTAGGTA	92810
QY	627	-----	627
Db	92811	AATTACTTAATCTCTGAGCTTCAGTTTTCTTATCTGTAAAGCAAGGTTGTAGTATTTT	92870
QY	627	-----	627
Db	92871	CAC TATTTAGAGGGTTGCAAGGATAAAGTATTAATAACTACACAAAGTGCA TAGGACAG	92930

QY	627	-----	627	QY	741	-----	741
Db	92931	TGTTTGACACATTGAAGCTCTCACTCTCAAGTGGGTCTCCCTACTGTATATGGTGGCAT	92990	Db	94009	CCTAGGATCAGGGTCTCTCTTTTAGTCCAGAGAGCCTCTGTCTATGAGAGAGCTTTGTT	94068
QY	627	-----	627	QY	741	-----	741
Db	92991	CCCTCCTTCTGGCCTCAGCAACCAACCCCACTGCCCCAGTGTGTGATCCTTCAGAA	93050	Db	94069	TGTTCACTGTCTTAGGCTGCCACCTTATTTTGCATTAACCATGAATAGTGTGATTTCCCA	94128
QY	627	-----	627	QY	741	-----	741
Db	93051	AGACTTCAAGAGCGCGATTTCTGTCCAGGTGCTATCAGCCTACTCCAGAGAGTGTCTCTT	93110	Db	94129	AAGTTTGAACAGAGGGAAACCTATATGTTGATCAGTAACCTCCCGACCCCAAGTTTGACTG	94188
QY	627	-----	627	QY	741	-----	741
Db	93111	GCTGACCCCTTTGCAATGCCAGCGGACACCCCTGCATTCCTGCACTCTGTGTGTTGTA	93170	Db	94189	TAAAAAAGTTTAAAAAACACATTTGTCTACATCACACTATCAGACTATCTCTCAAG	94248
QY	627	-----	627	QY	741	-----	741
Db	93171	GAACAGTGAGTTCACTCTAGACCAAGCCAGGACTGGCACTTAGAGTGGTGGCTGAGTG	93230	Db	94249	TCTAAGAAGAGTGAGTATCCCTGTGTGACCTTAACGGGAAACTGAGGCTCAGAGAGGAAAA	94308
QY	627	-----	627	QY	741	-----	741
Db	93231	GTGATGTGTGGGAGATGCTGGCAAAAGGCTCTCACTGAATGCCATCTTCTGTCTCTT	93290	Db	94309	GGAACCTTGTCCAAGGTCAACAGAGCCTAGGCTACAACTGGGACTCTCTGGTTTCCAGGCA	94368
QY	627	-----	627	QY	741	-----	741
Db	93291	CCTCATGCCCTTCTCCTGTCTCTATCCCTGGCCTCTCTCTCATCTGCAATCTCTGT	93350	Db	94369	AAACTTTTCCACTGCACACTACACTGCTGTGTGGCCCTGCAATTGGACACCTCAGTTGCCCA	94428
QY	628	-----	636	QY	741	-----	741
Db	93351	CCTTAATTTTCCATCTCTCTCTGCCCCCAGTCCCTTTTGGCCCTGCAGAGTTGAA	93410	Db	94429	TCATGATACTCCCTCACCCCTAACTTCCAGCCACATGGTTTACAGAGGGAGTAATCAA	94488
QY	636	sValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGln	656	QY	741	-----	741
Db	93411	GGTGAGGCAAGATGAGTCCCTGTGTGTGTCATGGCAGCCACCCCTCTCACCCACCCA	93470	Db	94489	CCAGAGCCATGAAGCTTGGGTTCTTTTTTTTCTTTTGAGACAGAGTCTCAGTCTGTG	94548
QY	656	nIleSerGlyTrpLysLeuTrpGlyValGlyThrGluGluAlaAspGlyAs	676	QY	741	-----	741
Db	93471	GATCTGTGCTACAACTATATTGGCGGAGGTGGGGCTGAGGAGAGGCCAATGGCGA	93530	Db	94549	CCCAGGCTGGAGTCAATGGTATGATCTCAGCTCACTGCAACCTCTGCTCTCGGTGTTCA	94608
QY	676	pArgProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLys	696	QY	741	-----	741
Db	93531	TGCGCTGCCAGAGGGGCCGTGGAGACCAAGCTTGGGATGTGGGGCTGTCCGCTCAAGAA	93590	Db	94609	AGCAATCTGTGTCTCGCTCGGCTCTTTAGCAGCTGGGATTTACAGGTGTGCACCATCTGG	94668
QY	696	sLysValLysGlnTrpGluLeuThrGlnLeu	706	QY	741	-----	741
Db	93591	GAAGTGAAGCAGTATGAGCTGACCCAGCT-AGGTGAGGAGGGCTGGGGGGAGGGAGG	93649	Db	94669	CTAATTTTGTATTTTATAGTAGAGACGGGGTTTACCATGTTGGCCAGGCTGGCCTTGA	94728
QY	706	-----	706	QY	741	-----	741
Db	93650	GAGGGATAGCTTTCAGTGGGAGCTGGGCAAGTTGTGTGAAGAAGTTAGAGATGGGATGC	93709	Db	94729	CTCCCGACCTCAGGTGATCCACCCATCTTGGCCTCCCAACGTGCTGGGATTACAGGTGTG	94788
QY	706	-----	706	QY	741	-----	741
Db	93710	TCTCTGTGAGCGGAGAGTGACCCCAAGATGTCCCCCTAGCCATTAGCCCTGTACCTTC	93769	Db	94789	AGCCACCACACCCAGCTGAGTTCTAAATTCAGTTTGTTCATTTCTTTCAGCAAAATAT	94848
QY	707	----ValProGlyArgProTrpGlyValLysLeuValAlaPheAsnLysHisGluAspGln	725	QY	741	-----	741
Db	93770	CCAGTCTCCCTGGCGGCTGTACGAGGTGAAGCTCGTGGCTTTTCAACAAACATGAGGATGG	93829	Db	94849	TTGCTGTAGTGTCTACTAAGTGCAGCTGCTTTGGGAATTTAGTGGTGGAACAACAAT	94908
QY	725	YTrAlaAlaValTrpLysGlyThrGluLysAlaProThrProAsp	741	QY	741	-----	741
Db	93830	CTATGACAGAGTGTGGAGGGCAAGACGAGAGAGCGCCGCCACCCAGG-TGAGGGCGGTG	93888	Db	94909	AAGGTTCTTGTCTCTTGGGGTTTATTTTATGTTTCTAAATAAACACTTATAAATATAG	94968
QY	741	-----	741	QY	741	-----	741
Db	93889	GGGGAAGAAGCGGGAGGGCTCAGGTGAGGGCTCTGCTAGCGGCTTCTCTCTTTTCC	93948	Db	94969	TGTTTATTATATTCAATTTTCTATCTATTTTAAAAATATTAAATAATTTAATCCTCAT	95028
QY	741	-----	741	QY	741	-----	741
Db	93949	TTTGGGCCAGCTGGTGAGCAGGTACTTTTAGTAGCCTCAGCTCATGCCCCAACCCATTCT	94008	Db	95029	ACAACCCCATGAGGTAGGTACTATCATTTAGAAATAGTGAATGGAGACACAGAGAGGTAA	95088
QY	741	-----	741	QY	741	-----	741
Db	93949	TTTGGGCCAGCTGGTGAGCAGGTACTTTTAGTAGCCTCAGCTCATGCCCCAACCCATTCT	94008	QY	741	-----	741

Db	95089	TAATTTCTCCAGGTCA	CACAGCTAGTAAGTGCAGAGCTGGAATTTGAACCAAGGCAGC	95148
Qy	741	-----	-----	741
Db	95149	CTGGTCCAGAGTCTGGTTCCATGGTTATTTCATTGCTGTGATTGGCTCTCTCTTTA	95208	
Qy	741	-----	-----	741
Db	95209	TTACGTTCTTACAGACTAGTGTTCACAGTCAAAATCTTGAAGCCTCAATTG	95268	
Qy	741	-----	-----	741
Db	95269	CCTTATCTGTGAAGTAGAATGATGCCACTCAACTCAGAGTCAAAAGGAGTGGTCAAA	95328	
Qy	741	-----	-----	741
Db	95329	GAACAGCCTTGGGAACGTGTAAACCACTATACATATTTAAGGATGAGTATCCTCCCTCC	95388	
Qy	742	-----	-----	751
Db	95389	TTCTCTACACTCTAGTTTTCACCTTTCAGACATGCTATCCAGAGGGGACCACTCCCTGCC	95448	
Qy	751	oProAlaHisValHisAlaGluSerAsnSerSerThrSerIleTrpLeuArgTrpIysLy	771	
Db	95449	TCCAGCCCACTCCATCGGAATCAAAACAGCTCCACATCCATCTGGCTTCGGTGGAAAA	95508	
Qy	771	sProAspPheThrThrValIleValAsnTyrThrValArgPheGlyProTrpGlyLe	791	
Db	95509	GCCAGATTTACCAAGTCAAGATTGTCAACTACACTGTGCGCTTCAGCCCTGGGGCT	95568	
Qy	791	uArgAsnAlaSerLeuValThrTyrTyrThrSer-	802	
Db	95569	CAGGAATGCCCTCCCTGTGTCACCTATTACACAG- GTGGGGACTAATGCCTTTACTAATGT	95627	
Qy	802	-----	-----	802
Db	95628	AAAGAAGGAGGAAGATAGTGGGGGTGTGTGGGAAGGCAGAGGTGTGATCAGG	95687	
Qy	802	-----	-----	802
Db	95688	GAAGGGCCCTGGGATTGAGAAATGTCATGTGCTCACCTGTTACGCTGAGGCCACACAG	95747	
Qy	802	-----	-----	802
Db	95748	AGGGGTGCACAGGGTCCAGTCTTGCCTGCTGACCAAGATGGTAGCAAGAGGCCATCC	95807	
Qy	802	-----	-----	802
Db	95808	AATCCCGTCCCTCACAACAGAGCCTTGTCTCAAACTCCGGAGAATGTGACACAGGACA	95867	
Qy	802	-----	-----	802
Db	95868	GACCTGTTTCTCAACCTCAGAGGCTCCAGTGTGATGTGTAAGAGAACCCCACTG	95927	
Qy	802	-----	-----	802
Db	95928	TCCTCAGAGTGTCTCCCTCTACCTCTGTAGATGGGGAGACAGAGCCTGAGAGCCTGAGG	95987	
Qy	802	-----	-----	802
Db	95988	TTTGGCGAGGAAGTCGATATAAAAGCAACCGATATATACAACTGCCATGCAAGGCAAAA	96047	
Qy	802	-----	-----	802
Db	96048	TTAAGCCAAAGAAAGATGATCGGAGATGATTCATAGTGGGAGAAATAACTTGGATTAA	96107	
Qy	802	-----	-----	802
Db	96108	TCCAGGAAGGTTTCTCGAGGAAGGGGTCTGGTTTCGAGATTGCACTTGGAGCAGATAC	96167	
Qy	802	-----	-----	802
Db	96168	ATCCAGACTTCTCTAGGTCCCTGGGTGTGTAGTCTGCCTCCAGGCCAGGAGCAGCT	96227	
Qy	802	-----	-----	802
Db	96228	TCCCTTCCAGAGCTAGAGCTGGGGCTGGGGCTGACTGCAGGGGAGTAGATTTGGAGGAG	96287	
Qy	803	-----	-----	803
Db	96288	TCTGCCCTGTGGAGAGAGTGTGCTGAGAGCACTAGTCACGTTTCTTGCCTTAGCAGTTT	96347	
Qy	803	rglyGluAspIleLeuIleGlyGlyLeuLysProPheThrLysTyrGlyPheAlaValG	823	
Db	96348	TGAGAAGACATCTCATTTGGCGCTTGAAGCCATTCACCAATACAGTTTTCAGTGCA	96407	
Qy	823	nSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThrLeuPr	843	
Db	96408	GTCTCAGCGGTGGACATGATGGGCTTTGGGCTCTGTGTGTGAGGGCTCCACCTGCC	96467	
Qy	843	oAspArg-	845	
Db	96468	TGACCGTAGTGTCTCCAGCCCTTGGCCACCAACCATCTTTTGGCCAACTCTCTCTTG	96527	
Qy	845	-----	-----	845
Db	96528	CCCCAGCAGCAAGAAACATCTTCAGAACCCCAAGGTTTGGGCGAGAGCTGGGGGCAAG	96587	
Qy	845	-----	-----	845
Db	96588	AGTTTCTGGGAGGGGAGGAATGTTAAGTCCCTCCACACACACATACATGTTTTCTAGAA	96647	
Qy	845	-----	-----	845
Db	96648	CCCCAACTAGAACACAGACCTCATCTTGTCTTGTGTGCTTGTTCATGAGGAGGAAGAA	96707	
Qy	845	-----	-----	845
Db	96708	GGGAGTCAATCATCATGCTCTGTCCCCAGCTCTGACATCTCTCCCATCTCTCTCTCC	96767	
Qy	846	-----	-----	861
Db	96768	CGCACCCAGGGCCCTCC-ACACCCCATCTCGACCTGCGACTGAGCCCCCTGACCCGTC	96826	
Qy	861	rThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGlyTrLe	881	
Db	96827	CACGGTTGGCTGCACCTGGTGCCTCCCAAGAGCCCAAGGGGAGATCGTGGAGTATCT	96886	
Qy	881	uIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrTrGl	901	
Db	96887	GATCTCTGTACAGCAGCAACACACAGCAGCTTGAGCACCACTGGACCTTGTCTCACCAGCA	96946	
Qy	901	u-	901	
Db	96947	GGGTGAGGGCTCTGTGCCCCAGGGCATCTGCTCAAGAGCTTCCAGGCTCTCTCCGCTTCC	97005	
Qy	901	-----	-----	901
Db	97006	AGCAGTGAGAAACAGTGTGGGGTTTGTTCAGGGGCTACTGGGAGAGCAAGCAGAGCT	97065	
Qy	901	-----	-----	901
Db	97066	GCTGCTTAAGTGGAGAGCAGACTCTCCAGTCCCTCTCCGCCACCTCTGCAATCCACCC	97125	
Qy	902	-----	-----	907
Db	97126	ACCCCATCCACCATCCCTGCTCTCCCTGCTCCCTCCACCCCGGAGAAACATCTTCAGTGC	97185	
Qy	907	aglyValHisGlyLeuGluSerAspThrArgTyrPhePheLysMetGlyAlaArgThrGl	927	
Db	97186	TGAGTCCATGGCTTGAGAGCGACACTCGTACTTCTTCAAGATGGGGCGGCGACAGA	97245	
Qy	927	uValGlyProGlyProPheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSe	947	
Db	97246	GGTGGGACCTGGGCTTTCTCCCGCTTGCAGGATGTGATCACGCTCCAGGAGAGCTGTC	97305	


```

QY      1230 LysAlaLeuCyGProLeuThrValSerProSerLeuProArgAlaProValSerAla 1249
      |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      101682 AAATCCCTGCGCTTAGAGCCAGCCAGCCGCTGCCAGATCCCGGTCTCTCTCTCT 101741

RESULT 15
AC112162/c
LOCUS   AC112162.8 GI:21327375
DEFINITION Mus Musculus chromosome 9 BAC clone MGS1-296M6 ES cell line, complete sequence.
ACCESSION AC112162
VERSION   AC112162.8
KEYWORDS  HTG.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
REFERENCE
AUTHORS   Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenneck,W., Xi,C., Juels,P. and Kucherlapati,R.
TITLE     High Throughput Mouse Sequencing
JOURNAL   Unpublished
REFERENCE
AUTHORS   Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenneck,W., Xi,C., Juels,P. and Kucherlapati,R.
TITLE     Direct Submission
JOURNAL   Submitted (20-FEB-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
REFERENCE
AUTHORS   Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenneck,W., Xi,C., Juels,P. and Kucherlapati,R.
TITLE     Direct Submission
JOURNAL   Submitted (04-MAY-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
REFERENCE
AUTHORS   Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenneck,W., Xi,C., Juels,P. and Kucherlapati,R.
TITLE     Direct Submission
JOURNAL   Submitted (06-JUN-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
REFERENCE
AUTHORS   Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenneck,W., Xi,C., Juels,P. and Kucherlapati,R.
TITLE     Direct Submission
JOURNAL   Submitted (21-SEP-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
COMMENT  On Jun 6, 2002 this sequence version replaced gi:21321767.
-----Genome Center:
Center:  Harvard Partners Genome Center
Center Code:  HPGC
Web site:  http://www.hpcgg.org/Sequence/mouse.html
Contact:  hpgc@mcgill.mgh.harvard.edu

```

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features

listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550). Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double stranded sequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation as low coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as Low Quality.

-----Summary Statistics

Center project name: ALI
Sequencing vector: pSMART; AF399742
Chemistry: Dye-terminator Big Dye; 100t
Assembly program: Phrap version 0.990319
Contig length: 116580
Fraction of Phrap value < 40: 0
Error Rate in Consed: 0.00 per 10,000 bases
Number of N's in consensus: 0

----- Distribution of Quality < 40 Bases:

1000	5	10	15	20	25	30	35	40
900	Phrap Value range							
800	-----							
700	Location/Qualifiers							
600	1. 116580							
500	/organism="Mus musculus"							
400	/mol_type="genomic DNA"							
300	/strain="ES cell line"							
200	/db_xref="taxon:10090"							
100	/chromosome="9"							
0	/clones="MGS1-296M6"							

FEATURES

source	Location/Qualifiers
STS	235..369 /standard_name="D9Mit316"
repeat_region	258..303 /rpt_family=" (TG)n"
STS	612..1204 /standard_name="Nope"
repeat_region	complement(1584..1643) /rpt_family="B4A"
repeat_region	1710..1748 /rpt_family=" (TG)n"
repeat_region	complement(3793..3909)

repeat_region	/rpt_family="PB1D10"	4042..4191	Db	12486	GGGTCTTTGGAGCTACCGGGGCGAGATGTGGTCAATTTGGCAGCCCCCAGAGAACC	12427
repeat_region	/rpt_family="MIR"	4247..4435	Qy	253	ThrValValSerGlyGlnAenValMetGluCysValAlaSerAlaAspProThrPro	272
repeat_region	/rpt_family="B2_Mm2"	complement(4504..4623)	Db	12426	ACGGTAGTGTCTGGACAGAGGTAGTGTGATGGAGTGGCTCTGCTGACCCACCCCT	12367
repeat_region	/rpt_family="RSINE1"	4720..4831	Qy	273	PheValSerTrpValArgGln	279
repeat_region	/rpt_family="L2"	complement(4840..4879)	Db	12366	TTTTGTCTCTGGGTCCGACAGCGGTGAGTGTGAGGAAAGAGGAGGAGCTGGGGAG	12307
repeat_region	/rpt_family="MIR"	4938..5118	Qy	279	-----	279
repeat_region	/rpt_family="B2_Mm2"	6216..6250	Db	12306	TGGGCAAGATGGTACAGTATTTTCAGGGCTCCAAGTCAGGAGCTTCTGCCTAAAGATAGA	12247
repeat_region	/rpt_family="AT_rich"	complement(6251..6381)	Qy	279	-----	279
repeat_region	/rpt_family="RSINE1"	complement(6421..6566)	Db	12246	AGGAGAGATTGACTCTTTAGAGTTGCTTTTGACTTCCATAGCACATGCAAGCAAGCTC	12187
repeat_region	/rpt_family="B1_MM"	6593..6612	Qy	279	-----	279
repeat_region	/rpt_family="(T)n"	complement(6668..6790)	Db	12186	ACATGTACACACATTTCATGCAGTAATAAAGTCTAAAGAAAGTTGTATTAAAGCCAGTGG	12127
repeat_region	/rpt_family="B1_MM"	complement(6705..6820)	Qy	279	-----	279
repeat_region	/rpt_family="MIR2B"	complement(6883..6954)	Db	12126	TGGTGGCACACATCTCTTAATCCAGCATCTGGGGAGGCGAGGCGGGCTCTGAGTT	12067
repeat_region	/rpt_family="ID_B1"	complement(6982..7058)	Qy	279	-----	279
repeat_region	/rpt_family="ID4"	8050..8246	Db	12066	TGAGGCCAGCCTGGTCTACAGACCAAGTTCAGGGTAAACAGGGCTACTCAGAGGAACCC	12007
repeat_region	/rpt_family="B2_Mm2"	complement(8445..8719)	Qy	279	-----	279
repeat_region	/rpt_family="B4"	complement(9019..9119)	Db	12006	TACTTTGGAAATTTCTGGGGTGGGAGGAGAGAGACACAGAGAGACAGAGAGCGAG	11947
repeat_region	/rpt_family="MIR"	9285..9371	Qy	279	-----	279
repeat_region	/rpt_family="MIR"	complement(10773..10846)	Db	11946	GGACAGAGACAGAGACAGGACAGACACACACACACACACACACACACACACAC	11887
repeat_region	/rpt_family="PB1D10"	10859..10956	Qy	279	-----	279
repeat_region	/rpt_family="B4"	complement(11043..11141)	Db	11886	ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGTATTAAATATCTGCTCCG	11827
repeat_region	/rpt_family="PB1D7"	11848..11884	Qy	279	-----	279
repeat_region	/rpt_family="(TC)n"	11884..11922	Db	11826	CCCACCAGAGACCATTTAGCCATCGTTTACCACCTCCCTTCTGCTTTTCCACCTAGGA	11767
repeat_region	/rpt_family="(TG)n"	11930..11966	Qy	279	-----	279
repeat_region	/rpt_family="(TC)n"	complement(11967..11994)	Db	11766	TCAGAGTAAGAGGAACTGGGATTCAACACGCGGCTGTGATGCAGCTTTTCAGGGATGT	11707
repeat_region	/rpt_family="B4"	complement(11995..12127)	Qy	279	-----	279
repeat_region	/rpt_family="B1_MM"	complement(12128..12143)	Db	11706	AGTGGGAACTCTGTTAGTTGGAGAGGGTGTAGAGGACAGTTCCAGCTCCCTCCAGCC	11647
repeat_region	/rpt_family="B4"	complement(12170..12252)	Qy	280	-----	293
repeat_region	/rpt_family="RSINE1"	12853..12993	Db	11646	TGCTCTCTCTGTCACAGATCGAAAGCTATCTCCAGGATGTCTATCTGCTGGGCGG	11587
repeat_region	/rpt_family="B1_MM"	complement(13288..13367)	Qy	293	gThrAsnLeuLeuIleAlaSerAlaGlnProArgHisSerGlyValTrpValCysArgAl	313
repeat_region	/rpt_family="MIR"		Db	11586	GACCAATCTACTCATCGCCAGCGCAGCTCGGCACTCTGGAGTCTATGCTGCGGAGC	11527
repeat_region			Qy	313	AsnLysProLeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeu	331
repeat_region			Db	11526	CAACAAGCCCCGACCGGTGATTTCCGCACTGGCGCTGCTGAGCTCCGAGTGTGTGTA	11467
repeat_region			Qy	331	-----	331
repeat_region			Db	11466	GGAAGGCTGCTGCTGGGGACGTTAAGGATGGTGTGTGTAGAGGAGAGAGGGCGGT	11407
repeat_region			Qy	331	-----	331

Alignment Scores:
 Pred. No.: 2,84e-90 Length: 116580
 Score: 2466.50 Matches: 837
 Percent Similarity: 20.68% Conservative: 19
 Best Local Similarity: 20.22% Mismatches: 41
 Query Match: 37.35% Indels: 3248
 DB: 10 Gaps: 20

US-09-754-997A-2 (1-1252) x AC112162 (1-116580)

Qy 233 GlySerLeuGluAlaThrArgGlyGlnAspValValIleValAlaProGluAsnThr 252

Db	11406	CCTAGAAGACTAAATGTGTGAGGGGGACCAATCGGAGGAGTCTGAAATTTTCTGATGGTGG	11347
Qy	331	-----	331
Db	11346	TGATTTGGAGGACTGGATTCCTGAGTGGGGGGGGGGGAGGGAATGGGGCGATTT	11287
Qy	331	-----	331
Db	11286	GAGACGGGGGTGGGGGGCGGATTTAAAGCCTGGTTTGGGTTCATGAGGAGTGAGAGTGC	11227
Qy	331	-----	331
Db	11226	AGGCGAGCAGGAGTGGGGTGTGAGTTATGGAAGGAGATAGTGAATGGTCTAGGGGGTGT	11167
Qy	331	-----	331
Db	11166	GTGAAGCCAGAAATCTGTAGTGGTAGCACTCAGAAAACCTGTGGCAGGAGCGTAGTTGG	11107
Qy	331	-----	331
Db	11106	ATACCAGCCCTAGCTACAGAGTTGAGACCAGCTACAGAGTGAAAGTCTATCTCAAAAAC	11047
Qy	331	-----	331
Db	11046	AAAAATTTAAACAATGCAATGTCTTAGGCTATTAAAGTCCCTATGAGTGTGTGCAGGCC	10987
Qy	331	-----	331
Db	10986	ATGGTGAAGTATTCGGGTTCACAGACTGAGTTGAGACAAGGTTTCTGTTCACCTTCATAC	10927
Qy	331	-----	331
Db	10926	TCCAGCTAGCTGGCCCGGGAAATCTCGGGGAATCTCCTGTCTCTCTCCTAGGATGC	10867
Qy	331	-----	331
Db	10866	TGGGATTAATCAGTTGGGTGTGATGCTGAGCGCGGAGATCAGAAATTCAGGGCCAACTG	10807
Qy	331	-----	331
Db	10806	GGCTATAAAGAGCTTTCTCAATACACAGAAAACTGGGGGACCAGGCTTGCTGAAGCT	10747
Qy	331	-----	331
Db	10746	GGAGTAACCTGCTGGGTATTAAGAAGCCACGGGGAGGAAGTCAAGTTACAGCCTAGAGC	10687
Qy	331	-----	331
Db	10686	TCAAAAGCTGGAATCTGTGGCTTGGGAGAAAGGAGGACTAGCCAGGAGGGTCCGAG	10627
Qy	331	-----	331
Db	10626	TCTAGTGGACAGAGCTGGGCCGTATGACTAGGGGAGAGGACCTCAGGGTGAGGCTTGG	10567
Qy	331	-----	331
Db	10566	GCGCGGTGTGGACACTTGGACAGGGGTAGGCCCTCTTGTGAGAGTCTCCCTCTTGCCC	10507
Qy	332	-----AlaIaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerTh	350
Db	10506	CTCAGCTGCCCCAGCCATCTCGCAGCGGCCCGAGGGCTCTCGCGAGCGGGGCCAGCAC	10447
Qy	350	rAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeuHisThrLeuHis	370
Db	10446	CGCGCGCTTGTGTGCGCGCTGCCGGGAGCCACGGCCCGCGCTGCATGGCTGCACGA	10387
Qy	370	pGlyIleProLeuArgProAsnGlyArgVallyValGlnGlyGlySerLeuVa	390
Db	10386	CGGGATCCCGTTCGACCCCATGGCGCGTCAAGGTGCAGGCGGTGGCGGACGTTGGT	10327
Qy	390	lIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAl	410
Db	10326	CATCACTCAGATCGGCTTCAGGACGCTGGCTACTACAGTGGTGTAGCAGAAAAACAGCGC	10267
Qy	410	aglyThrAlaCysAlaAlaProLeuAlaValValArgGluGlyLeuProSerAl	430
Db	10266	GGGAACTCCTGTGCGCTGCGCCCTGGCGGTAGTGTGCGAGGGGCTGCCAGGCGC	10207
Qy	430	aProThrArgValThrAlaThrProLeuSerSerSerValLeuValAlaTTPGluAr	450
Db	10206	CCGACTCGGGTCAAGCCACCGCGCTGAGCAGCTCCTCTGTGTGGTGGCTGGAGCGC	10147
Qy	450	gProGluLeuHisSerGluGlnIleLeGlyPheSerLeuHisTyrGlnIleAlaAArgG1	470
Db	10146	CCCTGAGTTGCACAGCGAGCAATCATTTGGCTTCTCTTCTTACACCAAAAGGCAGGGG	10087
Qy	470	yVal-----	471
Db	10086	TAT-GTCTACTTTGACCGCGGAATCCAGTTAGGCAAGTATGGATGATATAATAA	10028
Qy	471	-----	471
Db	10027	GATGAGAGGAAGGATGGGAGCAATACTGCTGGAAAGTGTATCTAGTATTAGAGAAATGA	9968
Qy	471	-----	471
Db	9967	GTGTACAGCACTCCTTGACCGAGAGAAATGCTTCTCTTTCCAGGGTTAAAGAGCCATCT	9908
Qy	471	-----	471
Db	9907	TCCACCCCGCACTAACTGACGAGCGGCTCTTCAATTAGCAGGGAGCTTCTCTG	9848
Qy	471	-----	471
Db	9847	GCCTATTATGAAATCAATGCAATAGTAGACTAGATGCGAGAGGCAATTTCTATCTCTCAC	9788
Qy	471	-----	471
Db	9787	CCTAGATTGTCTGTAACAATTTAGCTATACATAGACCAGTAGGATCCCGACAGACTGCC	9728
Qy	471	-----	471
Db	9727	AGCTAGTCAGACATGGAGTGGGCTGTGACAGTCCCTTTGTCCCCCGGTGTACCACCTC	9668
Qy	472	-----AspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThrGluLeuGlnVa	489
Db	9667	AGGAGTGGACAATGTGGAGTACCAGTTTGCAGTAAACAATGACACACAGAGCTGCAGT	9608
Qy	489	lArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyrSerGlnLeuG1	509
Db	9607	TCGGGACCTGGAAACCAACACGGATTATGAGTTCTACGTGTGGCTTACTCCAGCTGGG	9548
Qy	509	yAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAsp-----	524
Db	9547	GGCCAGCCGAACTCCAGCCCGCCCTGGTGCATACACTGGACGATGGTAGGACCTCTGA	9488
Qy	524	-----	524
Db	9487	ACCTGAGTGACCTGGGACCCAGAGACACTCCCACTCAGGGGACAGTGTGGCCCAAGGCC	9428
Qy	524	-----	524
Db	9427	TCTTCCCTTGTGTGATGTCTGTGTGCCCTATAAAGCAGTGGGCGAGAGGTTCTTCTTC	9368
Qy	524	-----	524
Db	9367	CCAGTGTAAATGGGGAACCTGGGGCTCAGAAAATGAAAATGACTTGGCCAGCTCAGCAG	9308
Qy	524	-----	524
Db	9307	GGCAGCTGGGACTTGAGCCCGAGCTCTGAGCCCCCATACCATGCGGCCCTTCTTCTTAACA	9248
Qy	524	-----	524
Db	9247	TACATTTACTACCCAGCCATTGGATATGAAAGCTTTCTGGCCCCACACTGTCTCACTAGCCTT	9188

QY	524	-----	524
Db	9187	GGGATAGAAAGGATGTCTCCGATCTCCAGCGTAGGGCTCCCTGAGAGGCTCCCTCCC	9128
QY	524	-----	524
Db	9127	TCCCTCCCTCAGATCCTGGGCTGGGCTCTCTCTGACTTCTCTGACTTGTCTGTGACCTCTGGGCAAGT	9068
QY	524	-----	524
Db	9067	CTCTGATCACTCTGACTGTGTCTTATCTTCTCTGTTGGTGGAGCGGAGCTCACGCCAAT	9008
QY	525	-----ValProSerAlaIaProGlnLeuThrLeuSerSerProbenProS	540
Db	9007	GTTCTTCTGTCTCAGTCCCGCAGCGCAGCACCCAGCTTACCTTGTCCAGCCCAACCCCT	8948
QY	540	erAspIleArgValAlaIaProLeuProLeuProSerSerLeuSerLeuGlnValLeuL	560
Db	8947	CGGATCAGGGTGCATGCTGCTGCTCCCTCCAGCTGAGCAATGGACAGGTGCTGA	8888
QY	560	ysrLysIleGluTyrGlyLeuGlyLys-----	569
Db	8887	AGTACAAGATAGAGTACGGTTTGGGGAAGGAAGGTGAGTGGGCCCGGGAGGTACAGAGA	8828
QY	569	-----	569
Db	8827	TGTCAGTGGGAATGGGGGTCTCCCTAGATAGAGGAGGTGTCTCAGTGGAAAGTGAGCCT	8768
QY	569	-----	569
Db	8767	AATTGGGCGATGGGATGAGAAAGCTCTGGAAAGAGCATTTGGGAGACTAGATAGCCCAAT	8708
QY	569	-----	569
Db	8707	GTGTAAAGTATGTGTGCAAGCCCAAGGCTGAGTTCAATCTCAGACTTCTATGTTAAAAA	8648
QY	569	-----	569
Db	8647	AAAAAAGCCTGGCACAGTGTGTAGTGTCTGTTGTAATCCAGAGCAGGGAGGTGGAG	8588
QY	569	-----	569
Db	8587	ACACGGGACACCTTGGAGTCACTGGGCTAAACAGTGTCCAGGCCAGCGAGAGTCC	8528
QY	569	-----	569
Db	8527	CTGTGCAGNAGAGAGGTGGACAATACCAGAGAAATGACAACAGAGTTGTCTATCTGACC	8468
QY	569	-----	569
Db	8467	TCCACATGTATGTATACACACACCTGCACATAGTGTGAACACTCACACTCATGGACACACA	8408
QY	569	-----	569
Db	8407	CGACACAGGGTTTCGGGCGAGTTGAGGAAAAGGGAAGCAAGAGGGTCTAGTAGCCTGC	8348
QY	569	-----	569
Db	8347	GCACAGATAAGAGGGTGTCTAGTGGGACCTTGGAGCTGTTCCTTTCTGTGAGGGAGCA	8288
QY	569	-----	569
Db	8287	ATTTACCCAGTGGATCTTGGAAACCTTCTAGTGTGTTGTTGTTGTTATTTTAAAGT	8228
QY	569	-----	569
Db	8227	GTAATTTATTTATGTATATAGTTCTCTATCTGCACGCGCCACCTGCATGCCAGAG	8168
QY	569	-----	569
Db	8167	AGGGAATCAGATCCAGGTACAGATGGTTGTGAGCCACAGTGTGTGCGAAATTGAAC	8108
QY	569	-----	569

Db	8107	TCAGGACCTCTGGAGAGCAGTCTAGTCTCTCTAACAATTGAGCCATCTCTCCAGCCCCAG	8048
QY	569	-----	569
Db	8047	AAACCTTCTGTTTACACACCGAGAGAAGTGTGGAGGCTCCAGAGTGCACATGAGACA	7988
QY	569	-----	569
Db	7987	TGGGCTGTCTCTACTGCGGAGATGATGCTGTCTGTGGGGAGGTGATGTGTGGTGGAAAGG	7928
QY	569	-----	569
Db	7927	AATGGAGAGTGAGGAAAAAGAAAGTGTGTAGGGTCAAGGGTAAAGCTAAAGAGAGAGGCAG	7868
QY	569	-----	569
Db	7867	GAGGCTGATGGGAGACAGCAGTCTAGGGCTCTCTGACACCGAGGACTGTCTTAGTGTCTTAG	7808
QY	569	-----	569
Db	7807	CCAGCCATCAAGCCTGTGTAAAGGTGTGTACTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	7748
QY	569	-----	569
Db	7747	AAAAGTTTCATGCTTGTGTTTCTCCATATGGAAACGAGGCGCAATGGGCTTTCAGCAGCGATA	7688
QY	569	-----	569
Db	7687	GGTAGAAGCCCGAGGTCAACTGTGGAGGATGAAACCTGCTCCGTGTGAGGGGTCAGTG	7628
QY	569	-----	569
Db	7627	CCCAGAGGACCTTGGGACCTGAAATGGAAAAGTCAATGTGTATAAAATTTAAAGATTGATG	7568
QY	570	-----GluaspGlnValPheS	575
Db	7567	GTCCCCCAATGAAGACACCCAGCAGCCTATCTTTTGTGCCAGAAATCAGGTTTCT	7508
QY	575	erThrGluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysV	595
Db	7507	CCACCGAGTCCCTGGAAATGAGACACACTTGTAACTCACTTCAGGCCAAACAAG	7448
QY	595	alTyrArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpM	615
Db	7447	TGTACCGAGTCCGGATTTTCACTGGCACTGGCGCTGGCTATGGAGTCCCTTCTCAGTGA	7388
QY	615	etGlnHisArgThrProGlyValHisAsnGlnSerHis-----	627
Db	7387	TGCAGCACAGGACACCTGTGTGTGCAACAACAGAGCCATGGTATGGACCTACACGGGCAGA	7328
QY	627	-----	627
Db	7327	CTGGACAGCGGGGACAGGTGCTCTGTGGCCTTAGATATTGCCCTGTATTATTGCCCTGTATT	7268
QY	627	-----	627
Db	7267	TGCCACCCCCCCCCCAACAACAACACTTGTATGACACCGGAAGGCTGGGTGTAGA	7208
QY	627	-----	627
Db	7207	GTGAAGGCTGGGCTCAAAATATCTCTTGTTCACCTCTGTGACTTTTGACAGCCACTCGCA	7148
QY	627	-----	627
Db	7147	TCTCCAGAGTTCGGTACTACCAAGGATGATAGCAGGACTAGTCAATTGTAAATTATACACA	7088
QY	627	-----	627
Db	7087	CAGTGGTCCGCATAGAACAGGAGCTCTTGGGTTGAGACCACGGCTCACTTGCACAGAGTG	7028
QY	627	-----	627

Db	7027	CTTGCCTAGCATGTACAAAGCACTAAATTTCAATCCCCAGGACCACATATAAACCCAGCATAC	6966
Qy	627	-----	627
Db	6967	CTCTGATCCTGGCAGTTCAAGGTTCTGCCTCACCTACTAGTAGTTGGGGACCAGCCTAG	6908
Qy	627	-----	627
Db	6907	GAGGTATGAAACTCTGTATCAAGAAAGGAAGGAAACCAATTAAGAGCTAGCTGGCTTTGC	6848
Qy	627	-----	627
Db	6847	CAAGAGCCCAAAACAGTAGAGTGTTHAATGGCTACAAAGTCTCCAGGCAGGTGCTAGTG	6788
Qy	627	-----	627
Db	6787	GAGGCTCATGCTTTAATCCAGCAGCTCGAGAGTGGAGGAGGCAGAGATTTCTGAATTTG	6728
Qy	627	-----	627
Db	6727	AAGACGCCTGGTCTACACCAAGTTACAGAGAAACCTGTCTCAAAAAACCAAGAAA	6668
Qy	627	-----	627
Db	6667	GA AAAAAGACTACAAAGTCAGTACCATATTCAGACATTAAGACTATAGCTTTAAAAA	6608
Qy	627	-----	627
Db	6607	AAAAA AAAAAAGCTGGATTATGGAATTATAAGNATGGGCGGGCGGTGGCGCAC	6548
Qy	627	-----	627
Db	6547	GCATGCTTTGATCCCACTACTCGGAGGAGCAGGAGCGGATTTCTGAGTTGAGCCT	6488
Qy	627	-----	627
Db	6487	GGTCTACAAAGTAGTTCCAGGACAGCCAGGCTACACAGAGAAACCTGTCTCGAAAAA	6428
Qy	627	-----	627
Db	6427	AGAAAAGAAAAGAAAAAANAATGATGAATTAANAATATACAAAGGCTGCCAAGGTG	6368
Qy	627	-----	627
Db	6367	CCCAGCTAATAAAGGGCTTGCTCCATTTCTGGGGATCTGACTTCCAATGTCTCCCCAG	6308
Qy	627	-----	627
Db	6307	AA TTTATATGATGATGATCTACAGCTTGGGCTCTGACCGCCACAGCAGCACAATA	6248
Qy	627	-----	627
Db	6247	AA TTAATAA TACTTTAAATTTTAAAAATACATATACAAATGTTTGAATGGGTCT	6188
Qy	627	-----	627
Db	6187	GACACAGCAACCTCTACTCTTGGGCTCGAGTGTGCTCGTGTGATGGTCCC	6128
Qy	627	-----	627
Db	6127	AAATCTCCCTCATGATCTCAGCAGCCCTTGACCATACTAACTTTTCAAGAGACCTGCA	6068
Qy	627	-----	627
Db	6067	GAAGCCCTCTGCCCTTGACGGGGGCTTGGCGAGACACGCCCTCTGTCTCTTTGTGC	6008
Qy	627	-----	627
Db	6007	GTGCCGGCAGACGGCCTGCGAGCTAA CATGCAGAGCTCAAAACCCCGGAAGTGC	5948
Qy	627	-----	627
Db	5947	AGAGTGAGCAGTTAGGAAAAGCCAGGCACAGCTCTATGCTGTGTCCGCCCTGCGCT	5888

Qy	627	-----	627
Db	5887	GCACCGGCTCACCTCCCTGACCACTTCTCACTCTCTCTTCCCTCTGTCGCCCATCATC	5828
Qy	628	-----ValProPheAlaProAlaGluLeuIysValArgAlaIysMetGlu	642
Db	5827	TTCCTCTGCCCCCAGTTCCTTTGGCCCTGCGAGATTGAAGGTGAGGCGCAAGATGGAG	5768
Qy	643	SerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTrpLysLeu	662
Db	5767	TCCCTGGTGTGTCATGGCAGCGCCCTCACCCACCCAGATCTCTGGATACAACTC	5708
Qy	663	TyrTrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProProGlyGlyArg	682
Db	5707	TACTGGAGAGAGGTGGGAAACAGAGAGAGGACAGATGGTGACCGCCCGCCAGGGGTCT	5648
Qy	683	GlyAspGlnAlaTrpAspValGlyProValArgLeuIysIysValIysGlnTrpGlu	702
Db	5647	GGAGATCAAGCTTGGGACGTGGGCGCGTGGCGCTGAAGAAGAAAGTGAAGCAGTATGAA	5588
Qy	703	LeuThrGlnLeu	706
Db	5587	CTGACCCAGTTAGTGTAGTAGGTCTGGGCTTGGGAGGTCATGGGCAAGCTAAGCACACG	5528
Qy	707	-----ValProGlyArgProTrpGluValIysLeuVal	717
Db	5527	GAGAGTGACACTCATGGCCCTTCCCCAGTCCCTGGCAGGCTGTACGAGGTGAAGCTCGTA	5468
Qy	718	AlaPheAsnLysHisGluAspGlyTrpAlaAlaValTrpIysGlyIysThrGluLysAla	737
Db	5467	GCTTTCAACAAACACGAGGACGGCTACGCTGCTGTGTGGAAAGGCGAAGACGGAGAAGCG	5408
Qy	738	ProThrProAsp	741
Db	5407	CCCACGCCAGG-TGAGGGGAGGAGCGGGGGTCTTGAGGCCTCTTTTCTCTTTTCTT	5349
Qy	741	-----	741
Db	5348	TATCAGCCAGCAAGCGTGTGCATGCCACAGCCACCGGCGCCACCTCTCTGGGCAGAAAT	5289
Qy	741	-----	741
Db	5288	CCTCCTGCTCTAGACATGCCCTCATTTCTCAAGCTAAGCAGAGCGAAGCGAGATGG	5229
Qy	741	-----	741
Db	5228	AAAGAGAAACTCCGTGTCCGATGGGTGATTCCTGGGCTCAAGCTTTAACTTTAAACAGCT	5169
Qy	741	-----	741
Db	5168	TTTTAGATCGCATCCAGTATACCACATGGTCCGTGAGATTATCCCATTAATCTTTTTTT	5109
Qy	741	-----	741
Db	5108	TAAAGACTTACTTACTTAATTTATATATAGTACACTGTAGCTGTCTTCAGACACACC	5049
Qy	741	-----	741
Db	5048	AGAAGAGGGCACCGATCCCATTTACAGATGGTTGTGGCTTCTCGGAATTAAGCTCCAGAC	4989
Qy	741	-----	741
Db	4988	CTCTAGAAAAGCAGTCAGTGCTCTCAACCACTGAGCCATCTCTCCAGCCCCACCCATAA	4929
Qy	741	-----	741
Db	4928	TTCTTAAGCAGGGCAAGCACAAACACAGTGGTTGTCCACCTGACGATCCAGCCTGCACAG	4869
Qy	741	-----	741
Db	4868	TCAGAAGAAACCTGGGTCTTAATTCACGCGAGTTTCAGTTTTTTTTCATTGAGCAATATTTT	4809

QY 741 ----- 741
Db 4808 ACCTAAGTATCTACCAATGACAGCTCTGTTCTGGAACTTAGTAATGAACAACATGAAT 4749
QY 741 ----- 741
Db 4748 AAGGTATCTGCTCTCTGGAGCTTATTTTATGCTTATCTAATAAACAATTTGTCATCA 4689
QY 741 ----- 741
Db 4688 AATAGTATACACTGTCCCAAGCTTCTACCTGCTTTGTAAAGCAGCCTCCTAAGGACACT 4629
QY 741 ----- 741
Db 4628 CATGAGGGGACAGTACAGATGGTTTACGACGGAAGGTGATGCTACAGGCCACATGACCTG 4569
QY 741 ----- 741
Db 4568 AGTTGGTTTCTGCACCCCCCCCCCAAAATGTCCTCTGCTCCCCACATGACACAGGAG 4509
QY 741 ----- 741
Db 4508 CACACTCCCTCCCTCCCTCCACACAATAATGTGTGACTGAGAGATGACAGGATTAAG 4449
QY 741 ----- 741
Db 4448 AATTTAAATATGTTTAAAGATGTATTTATGTTATTTATATGTGTGACTACACTGTCACT 4389
QY 741 ----- 741
Db 4388 GTTGCTATCTTCAGACACAGCAGAGAGGTCCTAGATCCATATTATAGATGGTTGTGAGC 4329
QY 741 ----- 741
Db 4328 CACCATGTGGTTGCTGGAAATTTGAACTCAGGGACTCTGGAAGAGCAGTCAGTGTAAACC 4269
QY 741 ----- 741
Db 4268 GCTGAGCCCTCTCTTCAGCCCCCAGAAATTTTAAATATTTTAAAGAGTAATAATAAGAA 4209
QY 741 ----- 741
Db 4208 AAGAAGAAATGTCAAGATGTAGGTATTTGTCATCAGATCCTTTAAGAAATAGAAATGA 4149
QY 741 ----- 741
Db 4148 TGGCAGACAGAGGTTAAACAATTTCTCTAAGGCCACATACATAGCTAGGAAGTGGCAGCA 4089
QY 741 ----- 741
Db 4088 ATGGAGCTGGCTCCAGAGTCAGGGCATGCTTTCTCTGTCACTTCTCATCAGCCATGGG 4029
QY 741 ----- 741
Db 4028 ATATGTCACAGTCACTCTATGTGTGACCTTAAGCAAAAAAATCACTTCCATTTCTGGGA 3969
QY 741 ----- 741
Db 3968 GCCTCAGTGCAGTAGGATTAATGTACCCAACTCACAGGATCAAGAGAGGCCCGGTG 3909
QY 741 ----- 741
Db 3908 CCAGCTGTGGTATCACACAGCTGTAAACCCCTGTGTAAAGGACCAGAGGCACAGGATGG 3849
QY 741 ----- 741
Db 3848 AGAATTACAGGCCAGCTAGGCTACTACATAATGAGATTTCTGCTCAAAAAATAGAGGGG 3789
QY 741 ----- 741
Db 3788 GAAAGGCCAATGAACCTACTGAGACCATGTAGTCACCTACTCTCTGACACTTGGG 3729
QY 742 -----LeuProIleGlnArglyProProLeuProProAlaHisValHis 756

Db 3728 TTTCCCTCTTAGACCTGCTATCCAGAGGGGGCACCGCTGCTCTCTGCCCATGTCAC 3669
QY 757 AlaGluSerAsnSerThrSerIleTTPLeuArgTTPlyLySProAappPheThrThr 776
Db 3668 GCAGAGTCAACACAGCTCCACTTCCATTTGGCTTCGGTGGAGAGACGACACTTTACCACT 3609
QY 777 ValIysIleValAsnTyrThrValArgPheGlyProTTPGlyLeuArgAsnAlaSerLeu 796
Db 3608 GTCAGAGTTGTCAACTACACTGTAGGCTTCGGCCCCCTGGGGGCTCAGGAATGCTTCCCTG 3549
QY 797 ValThrTyrThr----- 801
Db 3548 GTCACCTACTATATACAGGTATCAGTGAGGGGGAGTGTAGCTGGGGTGGTAGACTAGCGG 3489
QY 801 ----- 801
Db 3488 GTTCAGCTGTGGTGAGGCCACACCCAGGATCTACACAGATGTCTTTTCCGTGACTGTCTGA 3429
QY 801 ----- 801
Db 3428 CCTGGGGGTAAATAGAGGGATGGGATTCCTCCCATCCCTCAGTGGGAGAAATGTGTGGTC 3369
QY 801 ----- 801
Db 3368 AGCCTTTGTGAGAGTGGATGCAAGACTAGAGGCTTCATCCGTTGGCTGGCACCCCTGCC 3309
QY 801 ----- 801
Db 3308 GGCCCATCATTCGGAAGATAGCTTAAAGAGCTGCTGTCTGTGATTTGGGAAGACACCTGTG 3249
QY 801 ----- 801
Db 3248 AAAGCCAAACAGAGAGAGAACTAAGCTGAGGATAGACACTTCAGGAGCGATTTGCTCCCC 3189
QY 801 ----- 801
Db 3188 AGGTGCAGAAATTACTTAGATTAATTCGGGAAAGCGTCCAGAGAGAGAGAGGGCTTGCTT 3129
QY 801 ----- 801
Db 3128 TGGAGATGAGACATGGAATAACATATGTCGGGCTTTCTGTGTTTGCCTGGCTGTGTGT 3069
QY 801 ----- 801
Db 3068 GGGAGGGAGGTGTGGGGCCCAACAGCAGCTCCCTCCAGGGTGGGGCTGGCTGGAGGC 3009
QY 801 ----- 801
Db 3008 AGGGCTTCTGTCTGGGTTTCTGCTGGCTTGAGGCCCTGTCTGAGGGCAGCTAGTCTTATC 2949
QY 802 -----SerSerGlyGluAspIleLeuIleGlyLeuLySProPheThrIlyS 817
Db 2948 CCTTCTCTGGAAGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCACTTTACCAAG 2889
QY 818 TyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValVal 837
Db 2888 TAGAGTTTGGGTACAGTCCCAAGAGTGGATATGATGGGCCCTTTGGCTCCGTCGTA 2829
QY 838 GluArgSerThrLeuProAspArg----- 845
Db 2828 GAAAGCTCCACCTGCTGACCG- TGAGTATCCACTGCGCGCTCCCACTGCTGGCCACCA 2770
QY 845 ----- 845
Db 2769 TTTTCTGGGTCTGAGGAAGGAAGAGAGGCTGTGCTCTCATCTCCCTTAGTTAGG 2710
QY 846 -----ProSerThrProProSer 851
Db 2709 GCTGTGCTGACTCTCCACCCCTTCCCTGTGCTGACGGGCTTCAACACCTCTCTCT 2650
QY 852 AspLeuArgLeuSerProLeuThrProSerThrValArgLeuHisTTPCysProProThr 871

```
Db 2649 GACCTGCGCTGAGCCCTCGACCATCCACCGTTTCGGTTACACTGGTGTCCGCCACG 2590
Qy 872 GluProAsnGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnPro 891
Db 2589 GAGCCCAATGGGATGATGGAGTATCTAAATCTCTACAGCAACACACACCCAGCC 2530
Qy 892 GluHisGlnTrpThrLeuLeuThrGlu----- 901
Db 2529 GAACACCACTGGACACTGCTCACACAGA - GGTAAAGGAGCGCCACCCCTCCTTCG 2471
Qy 901 ----- 901
Db 2470 CCTTTGTTCCAAAGGCTCTCGGATTTGGAGGCTGAAGCTACTGTCTGGGTGCAAGTTG 2411
Qy 902 -----Gly 902
Db 2410 GTTTGCCCATGGTGCTTTGGCTCTCGCCACCCCTGTCTCTCCCACTTCCCAACACAGA 2351
Qy 903 AsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheIysMet 922
Db 2350 AACATCTTCAGTGCAGAGTCCATGGCTTAGAGAGTGACACTCGGTATTTCTTCAAGATG 2291
Qy 923 GlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThrLeu 942
Db 2290 GGAAGCCCGCACAGAGGTGGGGCCCTTTTCCCGCTTGCAGGATGTGATTACTCTG 2231
Qy 943 GlnGluThrPheSer----- 947
Db 2230 CAAAGACATTTCTCAGGTACAGGAGGAGGAGAGAGGCCCTTTGGGACAACCTGGAGGCC 2171
Qy 947 ----- 947
Db 2170 CAAGACCCCTGGATTGGCTCCCAAGGCCCTGCTCTCTGGAATTACATTTTCTTTACTTT 2111
Qy 947 ----- 947
Db 2110 GGAACCTTAAGGTGTGGGGATGGAGAAGCAGTGACAGCTTGATAGAGGAGCTGTCATAG 2051
Qy 947 ----- 947
Db 2050 CTTTGAGGGGTCTTTTATTATTTTATTCTCTCTGTGTGTGTTTGGCAGAGTGCATATGCG 1991
Qy 947 ----- 947
Db 1990 AAGGTCAAGGACATCAGAGATTGGTTTGTGGGGATCTTGTCTTGGCTTGGAGGGTGG 1931
Qy 947 ----- 947
Db 1930 GTGAGCTCCTGAGTTCGTGAAGAAGTGGTCAATAGGCTGGATAGACACTCAACAAGGGCT 1871
Qy 947 ----- 947
Db 1870 CAATTAAGATACAAAATAGATCCAGGGCTCCTTATTGTAAATATTGAAGATGTGAAG 1811
Qy 947 ----- 947
Db 1810 ATGAGCCAGTGGGGCATTAGCCACATAGACTGAGGAGCATGAAGCCCTGAGTCCAAA 1751
Qy 947 ----- 947
Db 1750 TGCACACACACACACACACACACACACACACACACACCGTAAAACTTGGCTGTGGC 1691
Qy 947 ----- 947
Db 1690 TGCTGTAATCCTTCTGAGGATGCAGACAGGATATTTCTGGGGCTGAGAGGACTGGG 1631
Qy 947 ----- 947
Db 1630 GAAGTCAGTGATGTAGACCTCTGATCTCCACACTACATATACATATGCCCCTTAGATGACA 1571
Qy 947 ----- 947
Db 1570 TATACAGACATGACACACACATGTGTGCACCTAGAGCCACCTTGGTTATACATATATAC 1511
```

```
Qy 947 ----- 947
Db 1510 ACCATGAGAGCCCTTGTGTGCCAGGAGGAABACAACCTTCACTGAAGAGCTGATTCTATAGA 1451
Qy 947 ----- 947
Db 1450 GTCCAGTTTGTGAGGAGTGGCAGGTCTGTACAGAGGAGAACCGTGTGTGACTCAGT 1391
Qy 947 ----- 947
Db 1390 ATGAGGCCCAAGTAAGATCCCGTGGCTCTCTCTATTTCCAATCAGAAGGGAGTGAGCTTCCA 1331
Qy 947 ----- 947
Db 1330 AACAGGTGTGAACACAGGTGTGTACTGTCACACACTTTTGATCTCCGTCTCTGTATGAGC 1271
Qy 948 -----AspSerLeuAspValHis 953
Db 1270 ACTGCTTCTTCTGACCTTGGAGTCACACTGTTCTCCCGCCGACAGACTCTCTTGGATGTGCAC 1211
Qy 954 AlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuAlaCysMet 973
Db 1210 GCGGTACAGGGGATCATCGTGGTGTCTGCTGCGCTTCTCTGCTCTCTGCGCTGCAATG 1151
Qy 974 CysAlaGlyLeuArgGlnSerSer----- 981
Db 1150 TGTGCTGGCTTACGAGGAGCTCCCAAGGTGAACCGGGGAACTCACTGAGACCACTGAGT 1091
Qy 981 ----- 981
Db 1090 TACCAGAAAGGCTGCAGAGTGGACAAAGAGCGTGTCTCTCTGCTGCCCGGGAAGCC 1031
Qy 981 ----- 981
Db 1030 CCAGTATAGTAGGGCTATGAGAGCTGGAGTGAAGTTTCAGTGTCTACATGTGCCCCCATGTG 971
Qy 981 ----- 981
Db 970 AGGCCCTGGCGTCTCATCTCAGTAGCACAAAACAGTGGATAAATTTATTATTAGATTTC 911
Qy 981 ----- 981
Db 910 TGAGATGAAGGATTGAGGTGTGTTTCTCATGAGAATGTAGATAAACTTAGGGAAGCCCC 851
Qy 981 ----- 981
Db 850 AGGACTCCCCACCTCAGTAAACAGCGGGCTGAGTACTGTGGGCTTCCCTCATGCGCTTAA 791
Qy 981 ----- 981
Db 790 ATGGAAATGGCCCATCTCGCCATACAGGAAGGTGCCATCTGAGGGGTTCAGGCCGAGTG 731
Qy 981 ----- 981
Db 730 GACTTGTAFTCTCTCTTCTGTTGATTGACACACACAGGAAACCAAAATCCACTCTTG 671
Qy 982 -----HisArgGluAlaLeuProGlyLeuSerSerSer 992
Db 670 TCCGACCCCAATAACTTTGTTCTGTGATCAAGGGAGCCCTGCGCGATTTGCTCTCTCA 611
Qy 993 GlyThrProGlyAsnProAlaLeuTyrThrArgAlaArgLeuGlyProProSerValPro 1012
Db 610 GGCACCCCAAGGAACCCAGCGCTCTACACAGAGCTCGACTTTGGGCCCCCAGTGTCCCT 551
Qy 1013 AlaAlaHisGluLeuGluSerLeuValHisProArgProGlnAspTrpSerProProPro 1032
Db 550 GCTGCGCATGAGTTGGAGTCTCTGTCATCTCTGTCCTCGTCCCGAGGATTTGGTCCCAACCC 491
Qy 1033 SerAspValGluAspIysAlaGluValHisSerLeuMetGlyCysLeuValSerAspCys 1052
Db 490 TCAGATGTGGAAGACAAGGCTGAAGTACACAGCCTTATGGGTGGCAGTGTTCAGATTGC 431
```

```
Qy 1053 ArgGlyHisSerLysArgLysIle-SerTrpAla-----GlnAlaGlyGlyProAsnTr 1070
Db 430 CGGGGCCACTCCAAGAGAAAGGTGAGACTGAGTCTCTTCAGGTCAAGG----- 382
Qy 1070 pAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuTh 1090
Db 381 -----ACTCACAGAACCCATTCCTCAATCAAC 353
Qy 1090 rArgAlaLeuLeuProAlaGlyThrGlyGlnThrLeuLeuGlnAlaLeuValTy 1110
Db 352 A-----CCAGGGGGCACC-----AGAGCACAACTTCT 326
Qy 1110 rAspGly---IleLysSerAsnGlyArgLysLysProSerProAlaCysArgAsnGlnVa 1129
Db 325 -GACGGTGAGATGGCTCGCGCGCACACACACACACACACACACACACACACAC 267
Qy 1129 lGluAlaGluValIleValHisSerAsp-----PheGlyAlaSerLy 1143
Db 266 ACACACACAGGCACTGAACCATACAGAGAAAGCCTGTCCTACACCTGGGAGATCCCA 207
Qy 1143 sGlyCys-----ProAspLeuHisLeuGlnAspLeuGlu 1154
Db 206 GAGCTGTGTTTCGAGAGACCCCATCTTCTTCTTCACTTACCTAATATGGAG 155
```

Search completed: October 14, 2003, 08:42:16
Job time : 8508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 14, 2003, 06:11:05 ; Search time 3845 Seconds
(without alignments)
7913.967 Million cell updates/sec

Title: US-09-754-997A-2
Perfect score: 6604
Sequence: 1 MARADRGGLLVLTFCLLSA.....CPLTVSPSLPAPVSSAQVP 1252

Scoring table:
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Rgapop 6.0 , Rgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09754997/runat_14102003_061100_9718/app.query.fasta_1.1415
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09754997 @CGN 1 1 2865 @runat_14102003_061100_9718 -NCPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DSELP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2595	39.3	2626	11	AK079073	Mus muscu
2	1610.5	24.4	4185	11	AK083540	AK083540 Mus muscu
3	1326	20.1	900	12	B1185886	B1185886 UNL-P-FN-
4	1292	19.6	817	14	CA317106	CA317106 UI-M-FW0-
5	1252	19.0	740	13	BU701299	BU701299 UI-M-EX0-
6	1216	18.4	699	14	CD351162	CD351162 UI-M-PY0-
7	1204	18.2	713	14	CB524810	CB524810 UI-M-PY0-
8	1202	18.2	741	14	CD352400	CD352400 UI-M-GI0-
9	1174	17.8	719	14	CA751384	CA751384 UI-M-P00-
10	1098	16.6	631	14	CB518504	CB518504 UI-M-GH0-
11	1097	16.6	759	14	CB988160	CB988160 AGENCOURT
12	1082.5	16.4	703	14	BY712424	BY712424 BY712424
13	1047	15.9	591	13	BU057120	BU057120 UI-M-P00-
14	1046	15.8	2906	11	AK051027	AK051027 Mus muscu
15	1032	15.6	651	10	BB624536	BB624536 BB624536
16	1008	15.3	637	10	BB65894	BB65894 BB65894
17	1002	15.2	630	10	BB626002	BB626002 BB626002
18	1000.5	15.1	3230	11	AK011916	AK011916 Mus muscu
19	990.5	15.0	717	13	EX100607	EX100607 EX100607
20	977	14.8	553	14	CA316452	CA316452 UI-M-FW0-
21	971	14.7	632	10	BB649997	BB649997 BB649997
22	960	14.5	544	14	CA892088	CA892088 B0170A06-
23	945.5	14.3	1137	11	AK082162	AK082162 Mus muscu
24	927	14.0	578	14	CB607425	CB607425 AMGNNUC-N
25	910	13.8	595	12	BM256356	BM256356 518704 MA
26	882	13.4	504	14	CA876082	CA876082 K094SH09-
27	879	13.3	509	14	CA873325	CA873325 K0924D02-
28	803	12.2	455	14	CA883713	CA883713 B0105A08-
29	784	11.9	442	14	CA891757	CA891757 B0167G06-
30	778	11.8	435	14	CA892100	CA892100 B0170B03-
31	753	11.4	672	9	AW143015	AW143015 EST293310
32	725	11.0	747	13	BU382531	BU382531 603858107
33	708	10.7	438	14	CB751460	CB751460 AMGNNUC:N
34	700.5	10.6	836	14	CB994721	CB994721 AGENCOURT
35	689	10.4	524	9	AA947283	AA947283 ok20b04.s
36	687	10.4	564	28	AZ354558	AZ354558 IM0093B22
37	682	10.3	466	9	AI386698	AI386698 ta37B09.x
38	678	10.3	400	14	CB699715	CB699715 AMGNNUC:S
39	671.5	10.2	411	14	CA871247	CA871247 K0908H04-
40	667.5	10.1	3642	11	AK084609	AK084609 Mus muscu
41	638.5	9.7	485	9	AI741225	AI741225 wg26e07.x
42	601	9.1	353	14	CA886377	CA886377 B0126D12-
43	588	8.9	624	13	BU056331	BU056331 UI-M-P00-
44	575.5	8.7	500	12	B1182798	B1182798 UNL-P-FN-
45	564	8.5	939	13	BQ646105	BQ646105 AGENCOURT

ALIGNMENTS

RESULT 1
AK079073
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK079073
Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:933015G14 product:hypothetical protein,
full insert sequence.

AK079073.1 GI:26098261
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

2626 bp mRNA linear HTC 05-DEC-2002

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
PUBMED
AUTHORS

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2 Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuoka, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., De Bonaldo, M. P., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashina, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2626)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
PUBMED
AUTHORS

1 MetAlaArgAlaAAspThrGlyArgGlyLeuValLeuThrPheCysLeuLeuSerAla 20
15 ATGGCGCGGGCGGACACGGGCGGAGGGCTCTGGTGGTGGACCTTTGCTGCTGCTGCGCG 74
21 ArgGlyGluLeuProLeuProGlnGluThrThrVallyLeuSerCysAspGluGlyPro 40
75 CGCGGGAGCTGCCATTGCCCGAGAGACAATGTCAGCTGAGCTGTGATGAGGAGCC 134
41 LeuGlnValIleLeuGlyProGluGlnAlaValValLeuAspCysThrLeuGlyAlaThr 60
135 CTGCAAGTGATCTTGGGCGCTGAGCAGGCTGTGGTGGTGGACTGCACCTTGGGGGCTACA 194
61 AlaAlaGlyProProThrArgValThrTrpSerIysAspGlyAspThrValLeuGluHis 80
195 GCTGTGGCGCTCCGACCGAGGCTGACATGAGCAAGGATGAGGACACTGTATAGAGCAT 254
81 GluLeuLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu 100
255 GAGAACCTTGCACCTCTACCAATGGCTGCTCTGGTGGTGGTCTTCAACCTTAGGCAAGAA 314
101 AspSerAspAspGluGluAlaLeuArgIleTrpIysValThrGluGlySerTrpSerCys 120
315 GACAGCGATGATGAGAGAGCTCTTAGGATCTGGAAGGTCCTGAGGCGGACGCTATTCCTGT 374
121 LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLeuAlaThr 140
375 CTGGCCCAACAGCCCGCTAGGAGTGGTGGCCAGCCAGGTTGCTGTGTCTCAAGCTTGCCACA 434
141 LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg 160
435 CTCGAAGACTTCTCTCTGACCCCGAGTCCAGATTTGGAGGAGAGACGGGACGACGCGC 494
161 PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnVal 180

misc_feature 1..2626
/note="hypothetical protein (evidence: rncds)"

BASE COUNT 599 a 767 c 724 g 536 t

ORIGIN

Alignment Scores:
Pred. No.: 5,21e-165 Length: 2626
Score: 2595.00 Matches: 563
Percent Similarity: 84.66% Conservative: 22
Best Local Similarity: 81.48% Mismatches: 60
Query Match: 39.29% Indels: 46
DB: 11 Gaps: 11

US-09-754-997A-2 (1-1252) x AK079073 (1-2626)

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

PUBMED

AUTHORS

Db 2617 AAGAACCTGGAGAGCTACAGAGACTGCAGAGTCTTGGTACCCATGATGCCAGCAGC 2676

QY 1028 Trp 1028

Db 2677 TTC 2679

RESULT 3

LOCUS BI185886 900 bp mRNA linear EST 10-JUL-2001

DEFINITION UNL-P-FN-df-h-02-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone

ACCESSION BI185886

VERSION BI185886

KEYWORDS EST.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

AUTHORS 1 (bases 1 to 900)

TITLE Caetano, A.R., Johnson, R.K. and Pomp, D.

JOURNAL Generation and sequence characterization of a normalized cDNA library from swine ovarian follicles

MEDLINE Mamm. Genome 14 (1), 65-70 (2003)

PUBMED 22419904

COMMENT 12532269

Contact: Pomp, D

Department of Animal Science

University of Nebraska, Lincoln

Lincoln, NE 68583-0908, USA

Tel: 402 472 6416

Fax: 402 472 6362

Email: dpomp@unl.edu

Oligo-dT track not found, Not 1 site shown in beginning of sequence is likely internal to the message.

Seq primer: M13 -29

POLYA=No.

FEATURES

Source

1..900

/organism="Sus scrofa"

/mol_type="mRNA"

/strain="University of Nebraska, Lincoln Swine Selection Line"

/db_xref="taxon:9823"

/clone="UNL-P-FN-df-h-02-0-UNL"

/dev_stage="ADULT"

/lab_host="DH10B (Life Technologies)"

/clone_lib="UNL-P-FN"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; The UNL-P-FN library is a normalized library representing porcine ovarian follicles, ranging between 2.0 to 10.0 mm in diameter, collected during 7 days of the follicular phase of the pig estrous cycle. This library was derived from the library UNL-P-F2. The tag is a string of 5-6 nucleotides present between the Not 1 site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996.

BASE COUNT 149 a 304 c 291 g 150 t 6 others

ORIGIN

Alignment Scores:

Pred. No.: 8,27e-80 Length: 900

Score: 1326.00 Matches: 262

Percent Similarity: 97.12% Conservative: 8

Best Local Similarity: 94.24% Mismatches: 8

Query Match: 20.08% Indels: 1

DB: 12 Gaps: 0

US-09-754-997A-2 (1-1252) x BI185886 (1-900)

QY 247 AlalaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266

Db 7 GCTGCCCGCAGAACACCACTGGTGTCCGCGCCAGAGCGTGGTGTAGTGGAGTGTGTGCC 66

QY 267 SerAlaAspProThrProPheValSerTyrValArgGlnAspGlyIleYrsProIleSerThr 286

Db 67 TCGGCTGACCCCACTATCTGTCTGTGGTCCGACAGGACGGAGAGCTATTTCATACC 126

QY 287 AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHisSer 306

Db 127 GATGTCTATCTGTGGCGCGCACTGTCTATACACGCGCGCCANCCGG-CATCT 185

QY 307 GlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 326

Db 186 GCGGTCTACGTCTGCAGAGCCCAACAGCCCGTACGCGCGACTTTGCGCCGCCGCCGCC 245

QY 327 GluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThr 346

Db 246 GAGCTCCGCGTGTGGCGGCTCTCGCATCTCGCAGGCGCCGAGGCGCTATCTCGGACG 305

QY 347 ArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeuHis 366

Db 306 AGGCGGAGCACCAGCGCTTGTGTGCGCGCAACGGGTGAGCGCGCGCGCTTGGCGG 365

QY 367 TrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValIleValGlnGlyGly 386

Db 366 TGGCTGCACAAACGGCGCGCTGCGGCCCAACGGCGGTGTCAAGGTGACGGCGCGCGC 425

QY 387 GlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysValAla 406

Db 426 AGCAGCTGTGTCTATCAGCAGATCGGCTGCGAGATGCGCGCTACTACTACGATGCGT 485

QY 407 GluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgGluGly 426

Db 486 GAGAACCAACGGCGGCGCAGCGGTGCGCGCGCGCGCTGGCAGTAGTAGTGGCGGAGGT 545

QY 427 LeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeuVal 446

Db 546 CTGCCAGCGCGCCACAGCGGGTCAACGCGCCACGCGCGCTGAGCAGCTCCACTGTGCTGTG 605

QY 447 AlaTrpGluArgProGluLeuHisSerGluGlnIleGlyPheSerLeuHisTyrGln 466

Db 606 GCCTGGGAGCGCGCGGAGCTGCACAGCGAGCAGATCATCGGTTTCTCTCTTCACTAC 665

QY 467 LysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThrGlu 486

Db 666 AAGGCAAGCGGCGATGGACAGTGTGGATACCACTTGGCGTGGACAAATGACACCACTGAG 725

QY 487 LeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyrSer 506

Db 726 CTACAGGTTTCGGGACCTGGAGCCCAACACAGACTATGAGTTCTATGTGTGGCTACTCA 785

QY 507 GlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAsp 524

Db 786 CAGCTGGGGGGCGAGCGCGCAGCTCCACCCCGCTGGTGGTCCACACATTGGATGAT 839

RESULT 4

LOCUS CA317106 817 bp mRNA linear EST 26-NOV-2002

DEFINITION UI-M-FW0-cbm-d-21-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone

ACCESSION IMAGE: 6811462 5', mRNA sequence.

VERSION CA317106

KEYWORDS CA317106.1 GI:24535230

SOURCE EST.

ORGANISM Mus musculus (house mouse)

REFERENCE Mus musculus

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

JOURNAL 1 (bases 1 to 817)

COMMENT NIH-MGC http://mgi.nci.nih.gov/

Unpublished National Institutes of Health, Mammalian Gene Collection (MGC)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

```

source
  Location/Qualifiers
    1. .817
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6"
      /db_xref="taxon:10090"
      /clone="IMAGE: 6811462"
      /tissue_type="whole brain"
      /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
      /lab_host="DH10B (T1 phage resistant)"
      /clone_lib="NIH BMAP FMO"
      /notes="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
      Site 2: Not I; The library was constructed according
      Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
      1996. Denatured RNA was size fractionated on a 1% agarose
      gel. First strand cDNA synthesis was primed with oligo-dT
      primer containing a Not I site. Double strand cDNA was
      size selected according to mRNA size fraction, ligated
      with Ecor I adaptor, digested with NotI and then cloned
      directionally into pYX-Asc vector. The library tag
      sequence located between the Not I site and the polyA tail
      is AGCAGACAG. This library was created for the University
      of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
      Developing Mouse Nervous System', supported by National
      Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
      program coordinator."
BASE COUNT 180 a 254 c 148 t 1 others
ORIGIN
Alignment Scores:
  Pred. No.: 1.46e-77 Length: 817
  Score: 1292.00 Matches: 245
  Percent Similarity: 98.39% Conservative: 0
  Best Local Similarity: 98.39% Mismatches: 4
  Query Match: 19.56% Indels: 1
  DB: 14 Gaps: 0
US-09-754-997A-2 (1-1252) x CA317106 (1-817)
QY 1004 AlaArgLeuGlyProSerValProAlaAlaHisGluLeuGluSerLeuValHisPro 1023
Db 2 GCTCGACTTGGGCCCCCAGTGTCCTGCTGCCCATGAGTGGAGTCCCTCGTGCACTCCT 61
QY 1024 ArgProGlnAspTrpSerProProSerAspValGluAspLysAlaGluValHisSer 1043
Db 62 CGTCCCGAGATGGTCCCAACCCCTCAGATGTGGAAGCAAGCTGAAGTACACAGC 121
QY 1044 LeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIleSerTrpAla 1063
Db 122 CTTATGGGTGGCAGTGTTCAGATGGCGGGGCCACTCCCAAGAGAAAGATCTCTGGGCT 181
QY 1064 GlnAlaGlyClyProAspTrpAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySer 1083
Db 182 CAGGCGAGGGGAGCAACCTGGCAGGCTCTGCTGCTCCACCGGAGACCGGGAGACACTG 241
QY 1084 GlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeu 1103
Db 242 GGTCCAGGCGGAGCTGACCCCTGCTGCTGCTCCACCGGAGACCGGGAGACACTG 301
QY 1104 LeuLeuGlnAlaLeuValTrpAspGlyIleLysSerAsnGlyArgLysLysProSerPro 1123
Db 302 CTGCTGCAAGCCCTGGTATATGCGCATTAAGCAACGGGAGAAAGAGCGGTCCTCCCA 361

```

```

QY 1124 AlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGlyAlaSerLys 1143
Db 362 GCCTGCGAGGAATCAGGTGGAAGCTGAGGTCAATGTCCACTCCGACTTCGGTGCATCCAAA 421
QY 1144 GlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThrAlaGluThr 1163
Db 422 GGATGTCCTGAGACTCCACCTCAAGACCTGGAGCCAGAGAACCACTGATGTCAGAGACT 481
QY 1164 LeuProSerThrSerGlyAlaValAspLeuSerGlnGlnValAspTrpLeuGlyArgGlu 1183
Db 482 CTGCTCTCCACGCTCGGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTGGCGGAGGAG 541
QY 1184 LeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeuProGluAla 1203
Db 542 CTGGGAGGGTGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 601
QY 1204 AlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAlaIleGluAlaPro 1223
Db 602 GCGAGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
QY 1224 GlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSerLeuProArg 1243
Db 662 GGGAAAGCTGCGAGCCCAAGCCCTGTGTCTCTTACAGTCAAGCCCAAGCTTCCCAAGN- 720
QY 1244 AlaProValSerSerAlaGlnValPro 1252
Db 721 GCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 747
RESULT 5
BU701299 740 bp mRNA linear EST 09-OCT-2002
UI-N-EXO-bwy-1-18-0-UI.r1 NIH BMAP_EXO Mus musculus cDNA clone
IMAGE: 5704793 5', mRNA sequence.
BU701299
BU701299.1 GI:23624921
EST.
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 740)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5, Location/Qualifiers
1. .740
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6"
  /db_xref="taxon:10090"
  /clone="IMAGE: 5704793"
  /dev_stage="embryo 15.5 dpc"
  /tissue_type="whole brain"
  /lab_host="DH10B (T1 phage resistant)"
  /clone_lib="NIH BMAP EXO"
  /notes="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
  Site 2: Not I; The library was constructed according to
  Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
  1996. Denatured mRNA was size fractionated on a 1% agarose
  gel. First strand cDNA synthesis was primed with an
  oligo-dT primer containing a Not I site. Double stranded

```

cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCGGTGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT	173 a	225 c	200 g	142 t
ORIGIN				
Alignment Scores:	6.52e-75	Length:	740	
Pred. No.:	1252.00	Matches:	242	
Score:	97.98%	Conservative:	0	
Percent Similarity:	97.98%	Mismatches:	5	
Best Local Similarity:	18.96%	Indels:	3	
Query Match:	13	Gaps:	0	
DB:				
US-09-754-997A-2 (1-1252) x BU701299 (1-740)				
QY	654	ProThrGlnIleSerGlyTyrLysLeuTyrTrpGlyValGlyThrGluGluGluAla	673	
DB	2	CCACCCAGATCTCTGGATAAACTCTACTGGAGAGGTGGGAACAGAGGAGGCA	61	
QY	674	AspGlyAspArgProProGlyGlyValGlyAspGlnAlaTrpAspValGlyProValArg	693	
DB	62	GATGTGACCGCCCCAGGGGGTCTGGAGATCAAGCTTGGAGCTGGGGCCCGTGGG	121	
QY	694	LeuLysLysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGlu	713	
DB	122	CTGAAGAGAAGTGAAGCAGTATGAATGACCCAGTTAGTCCCTGGCAGGCTGTACAG	181	
QY	714	ValLysLeuValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLys	733	
DB	182	GTGAAGCTCGTAGCTTTCAACAAACACAGGAGCGCTACGCTGCTGTGTGGAGGGCAAG	241	
QY	734	ThrGluLysAlaProThrProAspLeuProIleGlnArgGlyProProLeuProProAla	753	
DB	242	ACGGAGAGGCGCCCAACGACACCTGCTATCCAGAGGGGGCCACCGCTGCTCTGCC	301	
QY	754	HisValHisAlaGluSerAsnSerThrSerIleTyrLeuArgTrpLysLysProAsp	773	
DB	302	CATGTCACGACAGTCAAAACAGCTCCACTTCCTATTTGGCTTCGGTGGAGAGCCAGAC	361	
QY	774	PheThrThrValLysIleValAsnTyrThrValArgPheGlyProTyrGlyLeuArgAsn	793	
DB	362	TTTACCACTGTCAAGATTGTCAACTACACTGTACGCTTCGGCCCTGGGGGCTCAGGAAT	421	
QY	794	AlaSerLeuValThrTyrTyrThrSerSerGlyGluAspIleLeuIleGlyGlyLeuLys	813	
DB	422	GCTTCCCTGGTCACTATACCGCTCTGGAGAAGA-ATTCTCATTTGGCGGCTCGAAA	480	
QY	814	ProPheThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyProPhe	833	
DB	481	CCATTACCAAGTACGAGTTTGGGTACAGTCCCAACGAGTGGATATGATGGGGCCCTTT	540	
QY	834	GlySerValValGluArgSerThrLeuProAspArgProSerThrProProSerAspLeu	853	
DB	541	GGCTCCGTCTGAGACGCTCCACCCCTGCTGACCGCCCTCAACACCTCTCTGACCTG	600	
QY	854	ArgLeuSerProLeuThrProSerThrValArgLeuHisTyrCysProProThrGluPro	873	
DB	601	CGCTTGAGCCCTTGACACATCCACCGTTCCGTTTACACTGGTGTCCCCCAGC-GAGCCC	659	
QY	874	AsnGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHis	893	
DB	-660	AATGGTGAGATTGTGGAGTATCTAATTTCTTACAGCAACAA-CACACCCAGCCCGAACAC	718	
QY	894	GlnTyrThrLeuLeuThrThr	900	
DB	719	CAGTGGACACTGCTCACCACA	739	

RESULT 6
CD351162
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CD351162 699 bp mRNA linear EST 29-MAY-2003
UI-M-FY0-cft-p-13-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE: 6952446 5', mRNA sequence.
CD351162
CD351162.1 GI:31142737
EST.
Mus musculus (house mouse)
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 699)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.
Location/Qualifiers
source

1. 699
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clones="IMAGE: 6852446"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pYX-Asc; Site_1: Ecor I; Site_2: Not I; the library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT
ORIGIN

135 a 221 c 202 g 139 t 2 others

Alignment Scores:

Pred. No.: 1.65e-72 Length: 699
Score: 1216.00 Matches: 227
Percent Similarity: 98.71% Conservative: 2
Best Local Similarity: 97.84% Mismatches: 3
Query Match: 18.41% Indels: 0
DB: 14 Gaps: 0

US-09-754-997A-2 (1-1252) x CD351162 (1-699)

QY 919 PhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAsp 938
DB 3 TTCTTCAGATGGAGGCCCGCACAGAGTGGGGCTGGCCCTTTCCCGCTTCAGGAT 62
QY 939 ValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIle 958

Db 542 GTCCCTGCTGCCCATGAGTTGGAGTCCCTCGTGCATCTCTGCNCAGGATTTGGTCCCA 601

Qy 1031 ProProSerAspValGluAAspLysAlaGluValHisSerLeuMetGlyGlySerValSer 1050
 |||||
 Db 602 CCACCTCAGATGTGGAGACAGAGGCTGAAGTACACAGCCTTATGGTGGCAGTGTTC 661
 |||||

Qy 1051 AspCysArgGlyHisSerLysArgLysLysSerTrpAlaGlnAlaGlyGly 1067
 |||||
 Db 662 GATTGCGGGGCCACTCCANGAGAAAGATCTCTGTGCTCANGCAGGGGGA 712
 |||||

RESULT 8
 CD352400 741 bp mRNA linear EST 29-MAY-2003
 LOCUS UI-M-GIO-cgj-p-07-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
 DEFINITION IMAGE: 6856280 5', mRNA sequence.

ACCESSION CD352400
 VERSION CD352400.1 GI:31143911
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 741)

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..741
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6856280"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dr
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGGAGACAG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

BASE COUNT 139 a 238 c 219 g 145 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,56e-71 Length: 741
 Score: 1202.00 Matches: 240
 Percent Similarity: 98.37% Conservative: 2
 Best Local Similarity: 97.56% Mismatches: 4
 Query Match: 18.20% Indels: 0
 DB: 14 Gaps: 0

US-09-754-997A-2 (1-1252) x CD352400 (1-741)

Qy 106 GluAlaLeuArgGileTrrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerPro 125
 |||||
 Db 2 GAAGCTCTTAGGATCTGGAGAGTCACTGAGGGGAGCTATTCTCTGTCTGGCCACAGCCG 61
 |||||

Qy 126 LeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSer 145
 |||||
 Db 62 CTAGGAGTGGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGCCACACACTCGAAGACTTCTCT 121
 |||||

Qy 146 LeuHisProGluSerGlnLleValGluGluLysGlyThrAlaArgPheGluCysHisThr 165
 |||||
 Db 122 CTGACCCCGGAGTCCAGATTGTGGAGAGACCGGACAGCACGCTTTGAATGCCACACC 181
 |||||

Qy 166 LysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnValThrValProGluGlu 185
 |||||
 Db 182 AAGGCGCTTCAGGCCCCCATCTTACTTGGGAAAGACACAGGTGACGCTGCTGAGAG 241
 |||||

Qy 186 ProArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAsp 205
 |||||
 Db 242 TCCCGGCTCATCACTCTTCCCAATGGCGTCTCTCCAGATCTTAGATGTCCAGGACAGTGAT 301
 |||||

Qy 206 AlaGlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAla 225
 |||||
 Db 302 GCAGGCTCTACCGCTCGGTGGCCACCAATTCAGCCGCCCAACGATTCAGCCAGGAGGCC 361
 |||||

Qy 226 SerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIle 245
 |||||
 Db 362 TCGTCACTGTGGCCCTCAGAGGCTCTTGGAGGCTACAGGGGGGAGGATGTGTCATT 421
 |||||

Qy 246 ValAlaAlaProGluAAsnThrValValSerGlyGlnAsnValValMetGluCysVal 265
 |||||
 Db 422 GTGGCAGCCCCAGAGNACACCAAGTGTCTGGACAGAGTGTAGTGTAGTGGTGGCTG 481
 |||||

Qy 266 AlaSerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSer 285
 |||||
 Db 482 GCCTCTGTGACCCCAACCCCTTTTGTGTCTGGGTCCGACAGGATGGAAACCTATCTCC 541
 |||||

Qy 286 ThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
 |||||
 Db 542 ACGGATGTCATCGTTCTGGCCGAGCAACCAAGCCCCCGACGCGTGTATTCGCCCACTCGCGC 601
 |||||

Qy 306 SerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
 |||||
 Db 602 TCTGGAGTCTATGTCGCGGAGCAACCAAGCCCCCGACGCGTGTATTCGCCCACTCGCGC 661
 |||||

Qy 326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
 |||||
 Db 662 GCTGAGCTCCGAGTGTCTGCTGCCCGCCAGCCATCTCGCAGGCGCCCGAGGCGCTCTCGCGG 721
 |||||

Qy 346 ThrArgAlaSerThrAla 351
 |||||
 Db 722 ACGCGGGCCAGACCGCG 739
 |||||

RESULT 9

CA751384

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA751384 719 bp mRNA linear EST 27-NOV-2002
 UI-M-F00-cdn-j-21-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
 IMAGE: 6830806 5', mRNA sequence.

CA751384
 CA751384.1 GI:25579257
 EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 719)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 151 a 183 c 180 g 117 t
ORIGIN

Alignment Scores:
Pred. No.: 1,37e-64 Length: 631
Score: 1098.00 Matches: 208
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 16.63% Indels: 0
DB: 14 Gaps: 0

US-09-754-997A-2 (1-1252) x CB518504 (1-631)

QY 470 GlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThrGluLeuGlnVal 489
DB 3 CGAGTGGACAATGCGAGTACCAGTTTGCAGTAACATGACACACAGAGCTGCAGGTT 62
QY 490 ArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyrSerGlnLeuGly 509
DB 63 CGGGACCTGGAAACCAACGCGATTATGAGTTCTACGTGGTGGCTTCTCCAGCTGGGG 122
QY 510 AlaSerArgThrSerProAlaLeuValHisThrLeuAspAspValProSerAlaAla 529
DB 123 GCAGCGGAACCTCCAGCCGAGCCCTGGTGCATACACTGACGATGTCCTCCAGCGGACGA 182
QY 530 ProGlnLeuThrLeuSerProAsnProSerAspIleArgValAlaTyrLeuProLeu 549
DB 183 CCCAGCTTACCTTGTCCAGCCCAACCCCTCGGACATCAGGTGGCATGGCTGCCCTCG 242
QY 550 ProSerSerLeuSerAsnGlyGlnValLeuLeuTyrTyrIleGluTyrGlyLeuGlyLys 569
DB 243 CCTCCAGCGCTGAGCAATGACAGGTGCTGAAGTACAAGATAGAGTACGTTTGGGGAAG 302
QY 570 GluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeuThrLeuAsnSer 589
DB 303 GAAGATCAGGTTTCTCCAGCGAGGTGCTTGGAAATGAGACACACTTACGTTAACTCA 362
QY 590 LeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGly 609
DB 363 CTTTCAGCCCAACAAGATGTACCGAGTCCGATTTACGTGGCACTGGCGTGGCTATGGA 422
QY 610 ValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGlnSerHisValPro 629
DB 423 GTCCCTTCTCAGTGGATGACGACAGGACACCTGGTGTGCACAAACAGAGCCATGTTCC 482
QY 630 PheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuValValSerTrpGln 649
DB 483 TTTGCCCTTCAGAAATGAAGGTGAGGGCAAGATGGAGTCCCTGGTGGTGTATGGCAG 542
QY 650 ProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrpGlyGluValGlyThr 669
DB 543 CCGCCCTCCACCCACCCAGATCTCTGGATACAACTCTACTGGAGAGAGGTGGGACAA 602
QY 670 GluGluAlaAspGlyAspArgPro 678
DB 603 GAGGAGAGGAGCAGATGGTACCGCCCC 629

RESULT 11

CB988160
LOCUS
DEFINITION
AGENCY 13892946 NIH MGC 147 Homo sapiens cDNA clone
IMAGE:30346067 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 759)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM384 row: p column: 12
High quality sequence stop: 606.
Location/Qualifiers

FEATURES
source

1..759
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30346067"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 147"
/note="Organ: placenta; Vector: pBluescriptR; Site 1:
all-XhoI; Site 2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.3 kb and normalized to 10^5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH MGC library."

BASE COUNT 161 a 229 c 237 g 131 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 2,02e-64 Length: 759
Score: 1097.00 Matches: 218
Percent Similarity: 94.17% Conservative: 8
Best Local Similarity: 90.83% Mismatches: 11
Query Match: 16.61% Indels: 3
DB: 14 Gaps: 1

US-09-754-997A-2 (1-1252) x CB988160 (1-759)

QY 378 GlyArgValLysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGln 397
DB 27 GGGCGGTCAAGTCCAGGCGCGGTGGCAGCTGGTTCATCACAGATCGGCTGCAG 86
QY 398 AspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAla 417
DB 87 GACGCGGTACTTACAGTGGTGGCTGAGAACAGCGGGGAATGCGTGGCTGGCGG 146
QY 418 ProLeuAlaValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThr 437
DB 147 TCGCTGGCGGTGGTGGCGAGGGGCTGCCAGCGCCCCCAGCGGGTCTACTGTCTAG 206
QY 438 ProLeuSerSerSerSerValLeuValAlaTyrGluArgProGluLeuHisSerGluGln 457
DB 207 CCACCTGAGCAGTCCGCTGTGTGGCTGGAGCGGCGCGAGATGCACAGCGAGCAG 266
QY 458 IleIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyrGln 477
DB 267 ATCATCGGCTTCTCTCCACTACCAAGAGGACGCGGGCATGCACAAATGTGAATACCA 326
QY 478 PheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAsp 497
DB 327 TTTGAGTGAACAACGACACCACTACAGGTTCGGAGCTCGGAACCCCAACACAGAT 386
QY 498 TyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerProAla 517

Db 387 TATGAGTTCTACGTGGTGGCTACTCCAGCTGGGAGCCAGCCGACCTCCACCCACCA 446

Qy 518 LeuValHisThrLeuAspValProSerAlaProGlnLeuThrLeuSerSerPro 537

Db 447 CTGGTGACACACTGGATGATGTCCTCCAGTCAGCACCACCCAGCTCTCCCTGTCAGCC 506

Qy 538 AsnProSerAspIleActGValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGln 557

Db 507 AACCTTCGGACATCAGAGGTGGCTGGCTGCTCCCTCCAGCTGAGCAATGGGGAG 566

Qy 558 ValLeuLysTrpLysIleGluTyTrpGlyLeuGlyLysGluAspGlnValPheSerThrGlu 577

Db 567 GTGGTGAAGTACAAGATAGATACGGTTTGGAAAGGAAGATCAGATTTCTCTACTAG 626

Qy 578 ValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyArg 597

Db 627 GTGGGAGGAATGACACACGCTATGCTGAACCTCGCTTACGCCAACAAGGTGTATCGA 686

Qy 598 ValArgIle-SerAlaGlyThrGly-----AlaGlyTyTrpGlyValProSerGlnTrp 614

Db 687 GTACGGATTTTCGGCTGGTACAGCAGCCCGGGCTTCNGGGCCCTCCAGGTGG 744

RESULT 12

LOCUS BY712424

DEFINITION BY712424 RIKEN full-length enriched, 10, 11 days embryo whole body

ACCESSION BY712424

VERSION BY712424.1 GI:27123733

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 703)

AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chochoia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perce, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yangisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kondo, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Iehi, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 2002, 563-573 (2002)

MEDLINE 22354583

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Iehi, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source 1..703

Location/Qualifiers

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="2810488H22"

/tissue_type="whole body"

/dev_stage="10, 11 days embryo"

/clone_lib="RIKEN full-length enriched, 10, 11 days embryo whole body"

BASE COUNT 164 a 185 c 218 g 134 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 1..75e-63 Length: 703

Score: 1082.50 Matches: 210

Percent Similarity: 86.53% Conservative: 2

Best Local Similarity: 85.71% Mismatches: 3

Query Match: 16.39% Indels: 30

DB: 14 Gaps: 1

US-09-754-997A-2 (1-1252) x BY712424 (1-703)

Qy 497 AspTyTrpGluPheTyTrpValValAlaTyTrpSerGlnLeuGlyAlaSerArgThrSerSerPro 516

Db 2 GATTATGAGTTCTACGTGGTGGCTACTCCAGCTGGGGCCAGCCGACCTCCAGCCCA 61

Qy 517 AlaLeuValHisThrLeuAspValProSerAlaProGlnLeuThrLeuSerSer 536

Db 62 GCCCTGGTGATACACTGGAGATGTCCCGAGCCAGACCCAGCTTACCTTGTTCAGC 121

Qy 537 ProAsnProSerAspIleActGValAlaTrpLeuProLeuProSerSerLeuSerAsnGly 556

Db 122 CCCAACCCCTCGACATGAGGGTGGCATGCTGCCCTCGCCCTCCAGCTGAGCAATGGA 181

Qy 557 GlnValLeuLysTrpLysIleGluTyTrpGlyLeuGlyLysGluAspGlnValPheSerThr 576

Db 182 CAGGTGCTGAAGTACAAGATAGATACGGTTTGGAAAGGAAGATCAGATTTCTCCACC 241

Qy 577 GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTy 596

Db 242 GAGGTGCTGGAATGAGACCAACTTACGTGTTAACTCCTTACGCGCAACAAGTGTAC 301

Qy 597 ArgValArgIleSerAlaGlyThrGlyAlaGlyTyTrpGlyValProSerGlnTrpMetGln 616

```

Db      302  CGAGTCCGAGATTTTCAGCTGCGCTGGCTGGCTATGAGTCCCTTCTCAGTGGATGCAG 361
QY      617  HsAtrqThrProGlyValHisGlnSerHis-ValProPheAlaProAlaGluLeuLy 636
Db      362  CACAGACACCTGCTGTGCAACACGAGCCATGATC----- 398
QY      636  sValaArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrG 656
Db      398  ----- 398
QY      656  nIleSerGlyTyrLysLeuTyrTrpGlyGluValGlyThrGluGluGluAlaAspGlyAs 676
Db      399  ----TCTGGATACAACTCTACTGGAGAGAGGTGGCAACAGAGGAGGAGGCGCATGCTCA 454
QY      676  pArgProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLy 696
Db      455  CGCCCCCAGGGGGTGGTGGAGATCAAGCTTGGAGCTGGGCCCGCTGGCGCTGAAGAA 514
QY      696  sLysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLe 716
Db      515  GAAAGTAGGACGATGATGACTGACCCAGTTAGTCCCTGGCAGGCTGTACAGGTTGAAGCT 574
QY      716  uValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLy 736
Db      575  CGTAGCTTTCAACANACAGAGACGCTACGCTGTGTGGAAGGCGAAGACGAGAA 634
QY      736  sAlaProThrPro 740
Db      635  GCGGCCCGCCGCA 647

RESULT 13
BU057120
LOCUS      591 bp mRNA linear EST 26-AUG-2002
DEFINITION UI-M-FOO-caa-e-21-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
VERSION IMAGE:6408572 5', mRNA sequence.
ACCESSION BU057120
KEYWORDS BU057120.1 GI:22497197
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapps-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyX-5.
Location/Qualifiers
1. .591
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6408572"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="PH108 (T1 phage resistant)"
/clone_lib="NIH_BMAP_F00"
/note="Organ: Brain; Vector: pyX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

```

FEATURES

source

```

1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pyX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TCAGAGAGCC. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
BASE COUNT 126 a 179 c 179 g 106 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3,55e-61 Length: 591
Score: 1047.00 Matches: 193
Percent Similarity: 98.98% Conservative: 1
Best Local Similarity: 98.47% Mismatches: 2
Query Match: 15.85% Indels: 0
DB: 13 Gaps: 0
US-09-754-997A-2 (1-1252) x BU057120 (1-591)
QY 1005 ArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeuValHisProArg 1024
Db 3 CGACTTGGGCCCCCAGTGTCTCTGCTGCCATGAGTTGGAGTCCCTCTGTGATCTCTCGT 62
QY 1025 ProGlnAspTrpSerProProSerAspValGluAspLysAlaGluValHisSerLeu 1044
Db 63 CCCCAGGATTTGGTCCCAACCCCTCAGATGTGGAGACACAGGCTGAAGTACACAGCCTT 122
QY 1045 MetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIleSerTrpAlaGln 1064
Db 123 ATGGGTGGCAGTGTTCAGATTGCGGGGCCACTCCAGAGAGAAGATCTCTCTGGGCTCAG 182
QY 1065 AlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySerGly 1084
Db 183 GCAGGGGGGCCAACTGGGCGAGGCTCTTGGGCGAGGCTGTGAGCTGCCCGGTAGTGGT 242
QY 1085 ProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeuLeu 1104
Db 243 CCAAGGCCCGGCTCTGACCGGTCTCTGCTGCTCCAGCGGGAACCGGGCAGACACTGCTG 302
QY 1105 LeuGlnAlaLeuValTyrAspGlyLysSerAsnGlyArgLysLysProSerProAla 1124
Db 303 CTGCAAGCCCTGGTATATGACGCCATNAAGAGCAACGGGAGAAAGACCGCTCCACGCC 362
QY 1125 CysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGlyAlaSerLysGly 1144
Db 363 TGCAGGATTCAGTGGAGCTGAGGTCAATTGTCCACTCCGACTTGGTGCATCCAAAGA 422
QY 1145 CysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThrAlaGluThrLeu 1164
Db 423 TGTCTCTGACCTCCACTCCAGACCTGGAGCCAGAGGAACCACTGACTGCAGAGACTCTG 482
QY 1165 ProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeuGlyArgGluLeu 1184
Db 483 CCTTCCAGCTCTGGAGCTGTGGATCTGTCTCAAGAGAGCANACTGGCTGGGCAAGGACTG 542
QY 1185 GlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
Db 543 GGAGGGGTGCCAACCAACCAACCCAGTGGGCGAGAGAGGCTCACCTGCTTGT 590
RESULT 14
AK051027
LOCUS
DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:D030056K15 product:putative neuronal cell
adhesion molecule, full insert sequence.
ACCESSION AK051027
VERSION AK051027.1 GI:26341679

```

KEYWORDS	HTC, CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aikawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsumoto, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, D., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.P., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	12127851
PUBMED	12127851
REFERENCE	
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source

1. 2906
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:D030056K15"
/db_xref="taxon:10090"
/clone="D030056K15"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="9 days embryo"
228..2669
note="unnamed protein product; putative
putative neuronal cell adhesion molecule (MGD|MG1:1202390,
GB|NM_008988, evidence: BLASTN, 100%, match=726)"
/codon_start=1
/protein_id="BAC34502.1"
/db_xref="GI:26341680"
/translation="MAEPTASPRRLPALRRPGPLPPLPPPPPLLLLLLLPLPAP
SLGHSALAFSEVPNDITANPGQIVLCKVEGTPPVQVQVNRKNGALPEGTHTL
LANGSLIHFRLEGGSPDEGDKVCAQNRGLVSRKARLQAATMSDPHHPQAV
TGEGVARQCQIHGPKPLITWKRVPIDTDDERTLVGPENLTLVHTATVLECVATG
PHCVASNIASVRVSHGARLVSGSGTJKEPTLVLPENLTLVHTATVLECVATG
PRFTVSRDLGRPIGEGVLGTGNIISDVTVHSGVYVCAANRPGTRVRAQ
RLVVAAPAFVHPQPSIRPAGTMTAQGEPPPHVTLKNGVGLGAGHVLKN
NNSTLSISGVGPEDEIYQCAENIAGSSQASARLVLMAGLPGPPNVRVAVSST
EVVWSPELAHTKEIIGYVHLHTRKADSPKLYQEAIVSKSTFOHLYDLPESTAYSP
YIKATYRGASLASVPTLSTLGEAPVPPPLSVLLGSSSLQLLWKPRLAQNGGF
KLFRPVSAFSPGILLPGTVSSNISQLDPSVTVYEVKLLAYNHQDGNATVRFVSL
KGASERTALTTPCDCKEDVTNHTSTTGVIGIHGVTICIFCVLFLFGQRGLVLC
KDVENLSPQGPGRSORDPGLALNGLSRGGQLSRDEKVPDAKLEQLPPTAGSAA
QPGSTPTDPAAPACBETLSMVLQGLFNLVAGRTTREATSPCAGPGVPAPQDIGPVP
LSEGTQTPPAVAAPQ"

CDS

note="unnamed protein product; putative
putative neuronal cell adhesion molecule (MGD|MG1:1202390,
GB|NM_008988, evidence: BLASTN, 100%, match=726)"
/codon_start=1
/protein_id="BAC34502.1"
/db_xref="GI:26341680"

BASE COUNT 582 a 937 c 835 g 552 t
ORIGIN

Alignment Scores:
Pred No.: 2,97e-60 Length: 2906
Score: 1046.00 Matches: 328
Percent Similarity: 38.49% Conservative: 135
Best Local Similarity: 27.27% Mismatches: 307
Query Match: 15.04% Indels: 434
DB: 11 Gaps: 34

US-09-754-997A-2 (1-1252) x AK051027 (1-2906)

Qy 24 LeuProLeuProGlnGluThr-----ThrVallyLeuSerCysAspGlu 38
342 TTGCGCTGCTGCGCCAGCGCTGGTCTGGGCCATTCTGCTGAACCTGCTTTCTGTG 401

Qy 39 GlyProLeuGlnValIleLeuGlyProGluGlnAlaValValLeuAspCysThrLeuGly 58
402 GAGCCAAATGATGATGATATTGCAACCTGGCGCAGCCCATAGTGTCTGGCTGCAAGTGGAG 461

Qy 59 AlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeu 78


```

Db      462  GGCAAC-----CTCCAGTGCAGAGTCTCTCTGGAGGAAGAAATGGAGCAGAGCTACCA 512
Qy      79   GluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSer----ProLeu 97
Db      513  GAGGCGACCCACACACCACTCTGCTGGCCAAATGGCTCTCTGATCACCACCACTTCAGCGCTG 572
Qy      98   GluGlnGluAspSerAspGluGluAlaLeuArgIleTrpLysValThrGluGlySer 117
Db      573  GAGCAGGAGGTAGCCCTTCAGAT-----GAGGGTGAC 605
Qy      118  TyrSerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValLys 137
Db      606  TATGAATGTGGCTCAGAACCGCTTGGGCTGCTGGTCTCAGTCGGAAGGCTCTCTCCAG 665
Qy      138  LeuAlaThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGly 157
Db      666  GCTGCAACTATGTCCGACTTCCACGTACACCCCGCGCTGCAGCGGTGAGGAGGTGGT 725
Qy      158  ThrAlaArgPheGluCysHisThrLysGlyLeuProAlaProIleThrTrpGluLys 177
Db      726  GTGGCCGCTTCCAGTGCCAAATCCATGGGCTTCCAGCCCTTGATCATTGGGAGAG 785
Qy      178  AspGlnValThrVal---ProGluGluProArgLeuIleThrLeuProLysTrpLeuLeu 196
Db      786  AACAGAGTTCCTCCATTGACAGGAGATGAGAGGTACACACTACTGCCCAAGGAGTCTCTG 845
Qy      197  GlnIleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSer 216
Db      846  CAGATAACAGGACTTCTGGGCTGAGGACAGCGGCATCTTCCACTGTGTGGCTCTCAACATT 905
Qy      217  AlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGlu 236
Db      906  GCTAGTGTGGGTGCACCCAGCGGCGCCAGGCTCACTGTGTCA-----GGCTCGGCTCA 959
Qy      237  AlaThrArgGlyGlnAspValValIleValAlaAlaProGluAsnThrThrValValSer 256
Db      960  GGGACCTTACAGGAGCCCACTTCTGTG---GGGCTGAGAACCTCACCCCTGACAGTG 1016
Qy      257  GlyGlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrp 276
Db      1017  CACCAGACTCGGCTGTGGAGTGTGTGCCACGGCAACCGCGCCCATTTGTGTCTCTGG 1076
Qy      277  ValArgGlnAspGlyLysProIleSerThrAspValIle---ValLeuGlyArgThrAsn 295
Db      1077  AGCGCTGTGATGGCCGCCCATTTGGGGTGAGGGGATCCAGGTGCTGGGCACAGGAAC 1136
Qy      296  LeuLeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLys 315
Db      1137  CTTATCATCTCAGATGTGACTGTCCAGCATTCAGGTGTCTATGTCTGTGCGACCAAG 1196
Qy      316  ProLeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAla 335
Db      1197  CCTGGCACCGGGTGAGAGGACGCGCGGTGGGTGGGTGTGTGTGTGTGTGTGTGTGTGT 1256
Qy      336  IleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCys 355
Db      1257  TTTGTTCAGCATCCCCAGTCCATCTCCAGCGCTGCTGGGACCAAGCATTTTCCACTGC 1316
Qy      356  ArgAlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArg 375
Db      1317  CAAGCGGAGGGGAGCCACCGCTCATGTGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1376
Qy      376  ProAsnGlyArgValLysValGlnGlyGlySerLeuValIleThrGlnIleGly 395
Db      1377  GCAGGAGGCCATGTGAGGCTCAAGAACAAATACAGCACACTGAGCATTTCTGGAGTCCGA 1436
Qy      396  LeuGlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAla 415
Db      1437  CCTGAAGTAGGGCCATCTACCAATGTGTGGCTGAGAACATTTGTGGCTCGTCGCAAGCC 1496
Qy      416  AlaAlaProLeuAlaValValAlaArgGluGlyLeuProSerAlaProThrArgValThr 435

```

```

Db      1497  AGCGCCAGGCTGACCGTGTCTGTGGGTGAGGGTTGCCAGGCGCCCCCAGCAATGTCCGA 1556
Qy      436  AlaThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSer 455
Db      1557  GCTGTCTGTGTCTCTCACTGAGTTGAGTGTCTCTGAGTGTGAGCTCTGCGCCACACC 1616
Qy      456  GluGlnIleIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGlu 475
Db      1617  AAAGAGATCATTTGGCTGTCTTGCACATCAGCAAGGCTGCCGACTCACCCCAAGCTAGAG 1676
Qy      476  TyrGlnPheAlaValAsnAsnAspThrThrGluLeuGlnValValArgAspLeuGluProAsn 495
Db      1677  TACCAGGAAGCTGTGACGAAGACACCTTTTCAGCACCTGTGTGAGAGACCTTGAGCCCTCC 1736
Qy      496  ThrAspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSer 515
Db      1737  ACTGTCTACAGCTTCTTACATCAAGCCTACACACCGAGGGGCGCCAGTTTACCTCTGTCTC 1796
Qy      516  ProAlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSer 535
Db      1797  CCCACCTGGCGCAGCACCTCTGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1853
Qy      536  SerProAsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerSerLeuSerAsn 555
Db      1854  TTGCTGGGCGAGCTCTCTTACACTGTCTATGGAAGCCCTGAGGCTGAGGCTGAGGCTG 1898
Qy      556  GlyGlnValLeuLysTyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSer 575
Db      1898  ----- 1898
Qy      576  ThrGluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysVal 595
Db      1898  ----- 1898
Qy      596  TyrArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMet 615
Db      1899  -----CGGCTGCGCCAA----- 1910
Qy      616  GlnHisArgThrProGlyValHisGlnSerHisValProPheAlaProAlaGluLeu 635
Db      1911  -----CACAA----- 1916
Qy      636  LysValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThr 655
Db      1916  ----- 1916
Qy      656  GlnIleSerGlyTyrLysLeuTyrTrpGlyValGlyThrGluGluAlaAspGly 675
Db      1917  -----GGAGGCTTCAAGCTGTTTACCGCCAGTCAGTCAACCTCC----- 1958
Qy      676  AspArgProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLys 695
Db      1959  -----TTCACTGGCCCATCTTGTCTCCCT 1982
Qy      696  LysLysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLys 715
Db      1983  GGGACTGTCTCTCTTACACCTCAGCAGCTTGCACCCAGCAGCTGTGACGAGGTGAAG 2042
Qy      716  LeuValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGlu 735
Db      2043  CTCTGGCTTACACCAACGACATGAGATGGA----- 2072
Qy      736  LysAlaProThrProAspLeuProIleGlnArgGlyProProLeuProProAlaHisVal 755
Db      2072  ----- 2072
Qy      756  HisAlaGluSerAsnSerThrSerIleTrpLeuArgTrpLysLysProAspPheThr 775
Db      2072  ----- 2072
Qy      776  ThrValLysIleValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSer 795
Db      2073  -----AAGGCCAGCTCCGATTT-----GTGTCTCTGAGGGGAGCTCT 2111

```

QY 796 LeuValThrTyrThrSerSerGlyGluAspIleLeuIleGlyGlyLeuValProPhe 815
 Db 2111 ----- 2111
 QY 816 ThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySer 835
 Db 2111 ----- 2111
 QY 836 ValValGluArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeu 855
 Db 2112 ----- 2112
 QY 856 SerProLeuThrProSerThrValArgLeuHisTyrCysProProThrGluProAsnGly 875
 Db 2151 GAAGATGTCACC ----- 2162
 QY 876 GluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyr 895
 Db 2163 ----- 2171
 QY 896 ThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAsp 915
 Db 2171 ----- 2171
 QY 916 ThrArgTyrPhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArg 935
 Db 2171 ----- 2171
 QY 936 LeuGlnAspValIleThrLeuGlnGlnThrPheSerAspSerLeuAspValHisAlaVal 955
 Db 2172 ----- 2177
 QY 956 ThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAla 975
 Db 2178 ACAGGCAATTGTTATCGCATCGCATCGGAGTCACCTGCATCATCTTCTCGCTCTCTT 2237
 QY 976 GlyLeuArgGlnSerSerHisArgGluAlaLeuProGlyLeuSerSerGlyThrPro 995
 Db 2238 ----- 2269
 QY 996 -----GlyAsnProAlaLeuTyrThrArgAlaArgLeuGlyProProSerVal 1011
 Db 2270 GTGTAAGGATGTGGAAACACGACTCTC ----- 2302
 QY 1012 ProAlaAlaHisGluLeuGluSerLeuValHisProArgProGlnAspTyrSerProPro 1031
 Db 2303 CCAGGGTCCCGGAGCCGAGGAGGACCTGTGTATCTCTGGCCCT--GAATGACTGAGCCG 2359
 QY 1032 -----ProSerAspValGluAspLysAlaGluValHisSerLeuMetGlyGly 1047
 Db 2360 AGGAAAGGGGGCCAGCTGAGCCGAGATGAAGCCGGTGGAGCCGACGAGGTGGACAC 2419
 QY 1048 SerValSerAspCysArgGlyHisSerLysArgLysIleSerTrpAlaGlnAla----- 1065
 Db 2420 GCTGTTTCCAC -----AGCTGGGTGAGTGCACGACCC 2452
 QY 1066 -----GlyGlyProAsnTyrAla 1071
 Db 2453 AGGCTCACACCCACGGATCCTGAGCCCTGTCTCATGTGAGGAGGACCCAGCTCTCCAT 2512
 QY 1072 GlySer----- 1073
 Db 2513 GGTTCAGCTTCAGGCTTCACTTGTGGCAGGAGGACACAGAGCCACGTCCTCCCTG 2572
 QY 1074 -----TrpAlaGlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuThr 1090
 Db 2573 TGCAGGCGCTGGGCC-----TGTCACGAGCACCACGAGATATAGCCCTGTGCCCT 2623
 QY 1091 ArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeuLeuLeuGlnAlaLeuValTyr 1110
 Db 2624 CAGTGAGGAGCAGACCCAGCGCTCGAGCAGTCGACGACCCCGAGTGCATCTGTGTCC 2683

QY 1111 AspGlyIleLysSerAsnGlyArgLysLysProSerProAlaCysArgAsnGlnValGlu 1130
 Db 2684 CGAGGGCGGTATCCCAT -----TTCAGGTCAA 2713
 QY 1131 AlaGluValIle---ValHisSerAspPheGly---AlaSerLysGlyCysProAspLeu 1148
 Db 2714 GCGAGGTTTCTCTCTGCGGATTTGGCGCTCCGACGTAAGA----- 2761
 QY 1149 HisLeuGlnAsp-LeuGluProGluGluProLeu-----ThrAlaGluThrLeuProSe 1166
 Db 2762 -----CAGGGTCTCAACACACAGGACCGTGGTCTCTACAGCAGGACGCTCTCCCTG 2815
 QY 1166 rThrSer 1168
 Db 2816 CTCGAGC 2822

RESULT 15

BB624536
 LOCUS BB624536
 DEFINITION musculus cDNA clone 8030498014 5', mRNA sequence.
 ACCESSION BB624536
 VERSION BB624536.1 GI:15398322
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 651)
 AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuura,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,P., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished

TITLE

JOURNAL
 COMMENT
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ishii,Y. and Hayashizaki,Y.
 Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Funct. Genomics 2 pre, L72-L86 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
 e mouse tissues.

FEATURES

Location/Qualifiers

source 1. .651
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="8030498014"
 /sex="male"
 /tissue_type="testis"
 /dev_stage="15 days embryo"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 15 days embryo male testis"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCCAGAGCTCTTTTCTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 185.0 Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATTCGAGTTAATTAAATTAATCCCTCCCTCCCTCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC 1. Cloning sites, 5' end: SalI; 3' end: BamHI"

BASE COUNT 130 a 199 c 204 g 118 t
 ORIGIN

Alignment Scores:
 Pred. No.: 4.13e-60 Length: 651
 Score: 1032.00 Matches: 199
 Percent Similarity: 98.52% Conservative: 1
 Best Local Similarity: 98.03% Mismatches: 3
 Query Match: 15.63% Indels: 0
 DB: 10 Gaps: 0

US-09-754-997A-2 (1-1252) x BB624536 (1-651)

Qy 1 MetAlaArgAlaAspThrGlyArgGlyLeuLeuValLeuThrPheCysLeuLeuSerAla 20
 |||||
 Db 41 ATGGCGGGCGGACACACGGCGCGGGCTCTGGTGTGACCTTCTGCTGTCTGCCG 100
 |||||
 Qy 21 ArgGlyGluLeuProLeuProGlnGluThrThrVallysLeuSerCysAspGluGlyPro 40
 |||||
 Db 101 CGCGGGAGCTGCCATTGCCCCAGGACAACTGTCAAGCTGAGCTGTGATGAGGACCC 160
 |||||
 Qy 41 LeuGlnValIleLeuGlyProGlnAlaValValLeuAspCysThrLeuGlyAlaThr 60
 |||||
 Db 161 CTGCAAGTGATCTGGGGCCCTGAGCAGGCTGTGGTGTGGACTTGCACCTTTGGGGGTACA 220
 |||||
 Qy 61 AlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGluHis 80
 |||||
 Db 221 GCTGCTGGGCTCTCCGACAGGGGTGACATGGAGCAGGATGGAGACACTGTACTAGAGCAT 280
 |||||
 Qy 81 GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu 100
 |||||
 Db 281 GAGAACCTGCACCTGCTACCCCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAA 340
 |||||
 Qy 101 AspSerAspAspGluGluAlaLeuArgIleTrpIleValThrGluGlySerTrpSerCys 120
 |||||
 Db 341 GACAGCGATGATGAGGAGAGCTTTAGGATCTGGAGGTCACCTGAGGGCAGCTATTCTCTGT 400
 |||||
 Qy 121 LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThr 140
 |||||
 Db 401 CTGGCCACAGCCCGCTAGAGTGGTGGCCAGCCAGGTTGTGTGTCAAGCTTGGCCACA 460
 |||||
 Qy 141 LeuGluAspPheSerLeuHieProGluSerGlnIleValGluGluAsnGlyThrAlaArg 160
 |||||
 Db 461 CTGGAAGACTTCTCTCTGACCCCGAGTCCCAGATTGTGGAGGAGAACGGGACAGCGC 520
 |||||

Search completed: October 14, 2003, 09:25:32
 Job time : 3888 secs

Qy 161 PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnVal 180
 |||||
 Db 521 TTTGAATGCCACACCAAGGGCCCTTCCAGCCCCCATCATTTAGTTGGGAAAAAGACCAAGGTG 580
 |||||
 Qy 181 ThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAsp 200
 |||||
 Db 581 ACCGTGCTTGAGGAGTCCCGGCTCATCACTCTTCCCAATGGCGTCTCCAGATCCTAGAT 640
 |||||
 Qy 201 ValGlnAsp 203
 |||||
 Db 641 GTCCAAGAC 649

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 14, 2003, 06:11:04 ; Search time 511 Seconds
(without alignments)
6613.887 Million cell updates/sec

Title: US-09-754-997A-2

Perfect score: 6604

Sequence: 1 MARADTGRGLLVLTFCLLSA.....CPLTVSFLPAPVSSAQVP 1252

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPFO_spool/US09754997/runat 14102003 061059 9696/app_query.fasta_1.1415
-DB=N_Geneseq_19Jun03 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blos62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09754997 @CNG 1 1 0 @runat 14102003 061059 9696 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DISBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	6604	100.0	6176	22	AAD10021	Mouse Nope (neighb
2	5746	87.0	3753	22	AAI67204	Nucleotide sequenc
3	5610	84.9	3741	24	ABK11101	DNA encoding human
4	4903	74.2	2796	22	ABK10022	Mouse Nope (neighb
5	4361	66.0	3450	24	ABZ11219	Human polynucleoti
6	3167	48.0	2223	22	AAI67203	Nucleotide sequenc
7	1747	26.5	3453	25	ABX71435	Human cDNA encodin
8	1746	26.4	3210	25	ABX71432	Human cDNA encodin
9	1746	26.4	3874	25	ABX71447	Human cDNA encodin
10	1691	25.6	2976	25	ABX71437	Human cDNA encodin
11	1691	25.6	3219	25	ABX71440	Human cDNA encodin
12	1472	22.3	825	22	AAD10023	Mouse Nope (neighb
13	1458	22.1	3330	23	AA576658	DNA encoding novel
14	1384	21.0	2715	25	ABX71434	Human cDNA encodin
15	1384	21.0	2958	25	ABX71436	Human cDNA encodin
16	1329	20.1	2481	25	ABX71439	Human cDNA encodin
17	1329	20.1	2724	25	ABX71441	Human cDNA encodin
18	1144	17.3	2139	25	ABX71442	Human cDNA encodin
19	1144	17.3	2382	25	ABX71445	Human cDNA encodin
20	1067	16.2	2700	24	AAD36777	Human immunoglobul
21	958	14.5	4650	23	AA570835	DNA encoding novel
22	955	14.5	5553	24	AB576362	DNA encoding human
23	953	14.4	4820	25	ABT17038	Human MP21 gene ne
24	951	14.4	5297	24	ABV94279	Breast carcinoma r
25	951	14.4	5297	25	ABX10384	DNA encoding prote
26	943	14.3	4344	22	AAC90961	Human UNC-40 nucle
27	943	14.3	4608	16	AAQ80196	Deleted in colorec
28	943	14.3	4608	20	AAT23431	Human DCC DNA. Ho
29	943	14.3	4608	25	ABT17039	Human MP21 gene DC
30	943	14.3	5186	12	AAQ12752	Deleted in Colorec
31	933	14.1	2695	23	AA576339	DNA encoding novel
32	909	13.8	1875	25	ABX71443	Human cDNA encodin
33	829	12.6	4248	22	AAC90960	C. elegans UNC-40
34	781	11.8	1644	25	ABX71444	Human cDNA encodin
35	781	11.8	1887	25	ABX71446	Human cDNA encodin
36	768	11.6	6279	23	ABU14245	Drosophila melanog
37	698	10.6	2450	23	AA576657	DNA encoding novel
38	667	10.1	7630	23	ABL30223	Drosophila melanog
39	648	9.8	6734	24	AB199344	Mouse ischaemic co
40	647	9.8	1143	25	ABX71433	Human cDNA encodin
41	638	9.7	6000	16	AAQ86478	Human PTP-OB. Hom
42	638	9.7	6000	18	AA785389	Human protein tyro
43	638	9.7	6000	20	AA506095	Human protein tyro
44	638	9.7	6000	22	AA502144	Human protein tyro
45	637	9.7	6374	24	AAD36781	Human immunoglobul

ALIGNMENTS

RESULT 1
AAD10021
ID AAD10021 standard; cDNA; 6176 BP.
XX AC
XX AAD10021;
XX AC
XX AAD10021 (first entry)
DT 12-SBP-2001
XX Mouse Nope (neighbour of punc ell) cDNA.

XX Mouse; Nope; neighbour of punc ell; cytostatic; neuroprotective; vaccine;
XX gene therapy; cerebroprotective; colonic cancer; mental retardation;
XX tumour suppressor; chromosome 9; transgenic animal; genetic disorder;
XX obesity; Bardet-Biedl syndrome; autosomal recessive disorder; retinitis;
XX polydactyly; pigmentosa; hypogonadism; hypertension; diabetes mellitus;
XX renal anomaly; cardiovascular anomaly; ss.

Db 481 TTTGAATGCCACACAGAGGCTTCCAGCCCCCATCAATTACTTGGGAAAGGACCGAGTG 540
Qy 181 ThrValProGluGluProArgLeuLeuThrLeuProLysTrpLeuLeuGlnIleLeuAsp 200
Db 541 ACCGTGCTGAGGAGCCCGGCTCATCTCTCCCAAGTGGCTCTCTCCAGATCTAGAT 600
Qy 201 ValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArg 220
Db 601 GTCCAGACAGTGTATGAGGCTCTTACCGCTGCGTGGCCACCAATTTCAGCCCGCCACGA 660
Qy 221 PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly 240
Db 661 TTCAGCCAGAGGCTCGCTCACTGTGGCCCTCAGAGGGCTTTTGGAGGCTTACCAGGGGG 720
Qy 241 GlnAspValValIleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnVal 260
Db 721 CAGGATGTGTTCATTGTGGGAGCCCGGAGAACACACACCGGTAGTGTCTGGACAGATGTA 780
Qy 261 ValMetGlnCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp 280
Db 781 GTGATGAGTGGCGGCTCTGTCTGACCCCACTTGTGTCTGGTCCGACAGGAT 840
Qy 281 GlyLysProIleSerThrAspValIleValIleValLeuGlyArgThrAsnLeuLeuIleAsr 300
Db 841 GGAAGGCTATCTCCAGGATGTCATGTCTGGGCGGACCAATCTACTCATGCGCCAGC 900
Qy 301 AlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAsp 320
Db 901 GCGAGGCTCGGCATCTCGAGTCTATGTCTGCGAGCCCAACAGCCCTCTACGCGTGAC 960
Qy 321 PheAlaThrAlaAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro 340
Db 961 TTGCCACTGCGGCTGTGAGTCCGAGTGTCTGCTGCCAGCCATCTCGCAGGCACCC 1020
Qy 341 GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu 360
Db 1021 GAGCGCTCTGCGGAGCGCGGCGAGCACCGCGCTTCTGTGTGCGGGGCTCGGGGAG 1080
Qy 361 ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyValGluVal 380
Db 1081 CCAGGCGCGGCTGCTGCTGCTGCACGAGGGATCCCGTGTGACCAATGCGCGGCTC 1140
Qy 381 LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly 400
Db 1141 AAGGTGAGGCGGTGGCGGAGCTTGTCTATCATCTCAGATCGCGCTCGAGGAGCTGGC 1200
Qy 401 TyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAla 420
Db 1201 TACTACAGTGGGTAGCAGAAAACAGCGCGGAACTGCTGTGCGGCTGCGCCCTGGCG 1260
Qy 421 ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer 440
Db 1261 GTAGTGTGGCGAGGGGCTGCCGAGCGCCCGGACTCGGGTCAAGCACCGCGCTGAGC 1320
Qy 441 SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGly 460
Db 1321 AGCTCTCTGTGTGGTGGCTGGGAGCGGCTGAGTTGCACGAGAGCAATCATTTGC 1380
Qy 461 PheSerLeuHisTyrGlnLysAlaArgValValAspAsnValGluTyrGlnPheAlaVal 480
Db 1381 TTCTCTCTTCACTACCAAAAGGCAAGGGGAGTGGACAATGTGGAGTACCAAGTTTGCAGTA 1440
Qy 481 AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPhe 500
Db 1441 AACAAATGACACACAGAGCTGCAAGTTCCGAGCTTGGAACTTGGAACTTATGAGTTC 1500
Qy 501 TyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHis 520
Db 1501 TACGTGTGTGCTACTCTCCAGCTGGGGGCGAGCGAACCCTCCAGCCAGCCCTGTGTGCAT 1560
Qy 521 ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer 540

Db 1561 ACATGGACGATGTCTCCAGCGGACGACCCAGCTTACCTTGTCTCAGCCCAACCCCTCG 1620
Qy 541 AspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLys 560
Db 1621 GACATCAGGTGGCATGGCTGCCCTCCCTCAGCTGAGCAATGACAGGTGCTGAAG 1680
Qy 561 TyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly 580
Db 1681 TACAAGATAGGTACGGTTTGGGGAAGGAAGATCAGGTTTCTCCACCGAGGTGCTCGGA 1740
Qy 581 AsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIle 600
Db 1741 AATGAGACACACTTACGTTAACTCACTTACGCCAAACAAAGTGTACCGAGTCCGGATT 1800
Qy 601 SerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrPro 620
Db 1801 TCAGCTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGTGTCAGCACAGCACCT 1860
Qy 621 GlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys 640
Db 1861 GGTGTGCACACAGAGCCATGTTCCCTTTGGCCCTTGCAGAAATTGAAGGTGAGGGCAAG 1920
Qy 641 MetGluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTyr 660
Db 1921 ATGGAGTCCCTGTGTGTGTCTATGGCAGCGCCCTCACCCACCCAGATCTCTGTGATC 1980
Qy 661 LysLeuTyrTrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProProGly 680
Db 1981 AACTCTACTGGGAGAGGTGGGAACAGAGGAGGAGGAGATGTCAGCCGCCCCCGAGG 2040
Qy 681 GlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGln 700
Db 2041 GGTGTGTGAGATCAAGCTTGGGACGTGGGCGCGCTGGAAGAAAGTGAAGCAG 2100
Qy 701 TyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsn 720
Db 2101 TATGAATCTGACCCAGTTAGTCCCTGCGAGCGCGTACGAGGTGAAGTCTGTAGCTTTCAAC 2160
Qy 721 LysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysValProThrPro 740
Db 2161 AACACAGAGGAGCGGTACCGTGTGTGTGGAGGGCAGACGAGGAAGGGCCGACGCCA 2220
Qy 741 AspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsn 760
Db 2221 GACCTGCCTATCCAGAGGGGCGCACGCTGCCTCTCTGCCATGTCTCCAGCAGAGTCAAAC 2280
Qy 761 SerSerThrSerIleTrpLeuArgTrpLysProAspPheThrThrValLysIleVal 780
Db 2281 AGCTCCACTTCCATTTGGCTTCGGTGGAAAGAACCCAGACTTTTACCACTGTCAAGATTGTC 2340
Qy 781 AsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyr 800
Db 2341 AACTACACTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCCTGTGTCACTACTAT 2400
Qy 801 ThrSerSerGlyGluAspIleLeuIleGlyLysLeuLysProPheThrLysTyrGluPhe 820
Db 2401 ACCAGCTCTGGAGAAGACATTTCTCATTTGGCGGCTCGAAACCAATTTACCAAGTACGAGTTT 2460
Qy 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840
Db 2461 GCGGTACAGTCCACCGAGTGGATATGGATGGGCCCTTTGGCTCCGTCGTAGAACGCTCC 2520
Qy 841 ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
Db 2521 ACCCTCGCTGACCGGCTTCAACACCTCTCTTCTGACCTGGCGCTGAGCCCTTGACACCA 2580
Qy 861 SerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyr 880
Db 2581 TCCACGCTTCGGTTACACTGGTGTCTCCCGCCACGAGGCCCAATGCTGAGATTGTGGAGTAT 2640
Qy 881 LeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThr 900
Db 2641 CTAATTTCTTACAGCAACAACACACACCGAGCCCGGAACACCAAGTGGACACTGTCTCACCA 2700


```
QY 901 GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
DB |||||
DB 2701 GAGGGAACATCTTCAGTGCAGAGGTCCATGGCTAGAGAGTGACACTCGGTATTCTTC 2760
QY 921 LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle 940
DB |||||
DB 2761 AAGATGGAGGCCGACACAGAGGTGGGGCTGGGCCCTTTCCGGCTTCAGGATGTGATT 2820
QY 941 ThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIleIleVal 960
DB |||||
DB 2821 ACTCTGCAAGACACATCTTCAGACTCTTGGATGTGCAGCGCTCACGGGATCATCTGTG 2880
QY 961 GlyValCysLeuGlyLeuLeuCysMetCysAlaGlyLeuArgGlnSer 980
DB |||||
DB 2881 GGTGTCTGGCTGGGCTTCTTGGCTCTGGCTGCATGTGTGTGGCTACGACAAAGC 2940
QY 981 SerHisArgGluAlaLeuProGlyLeuSerSerGlyThrProGlyAsnProAlaLeu 1000
DB |||||
DB 2941 TCCACAGGGAAGCCCTTCCCGGATTGTCTCTCAGGCACCCAGGAAACCCAGCGCTC 3000
QY 1001 TyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuSerLeu 1020
DB |||||
DB 3001 TACACAGAGCTCGGCTTGGGCTTCCAGTGTCTCTGCTGCCCATGAGTTGGAGTCCCTC 3060
QY 1021 ValHisProArgProGlnAspTrpSerProProSerAspValGluAspLysAlaGlu 1040
DB |||||
DB 3061 GTGCATCTCTGCTCCCGGATTGTCTCCACACCTCAGATGTGGAGACAAAGCTGAA 3120
QY 1041 ValHisSerLeuMetGlySerValSerAspCysArgGlyHisSerIleArgLysIle 1060
DB |||||
DB 3121 GTACACAGCTTATGGGTGGCAGTGTTCAGATTCCCGGGCCACTCCCAAGAGAAAGATC 3180
QY 1061 SerTrpAlaGluAlaGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuPro 1080
DB |||||
DB 3181 TCTTGGCTCAGCAGGGGACCAAACTGGGAGGCTCTCTGGGAGGCTGTGAGTGGCC 3240
QY 1081 GlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
DB |||||
DB 3241 CAGGGTAGTGTTCGAAGCGCGCTCTGACCCGTGTCTGTCTCTCCAGCGGGAACCGGG 3300
QY 1101 GlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLysLys 1120
DB |||||
DB 3301 CAGACACTGCTGCTGCAGCCCTGGTGTATGACGCATATAAGAGCAACGGGAGAAAGAG 3360
QY 1121 ProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGly 1140
DB |||||
DB 3361 CGGTCCCGAGCTGGAGGAATCAGGTGGAGCTGAGGTCTATTGTCTCCACTCCGACTTCGT 3420
QY 1141 AlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThr 1160
DB |||||
DB 3421 GCATCCAAAGGATGTCTGACTTCCACTCCAGACCTGGAGCCAGAGAACCACTGACT 3480
QY 1161 AlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeu 1180
DB |||||
DB 3481 GCAGAGACTCTGCTTCCAGCGCTGGAGCTGGATCTGTCTCAAGGAGCAGACTGGCTG 3540
QY 1181 GlyArgGluLeuGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
DB |||||
DB 3541 GGCAGGGAGCTGGGAGGGTGCACCAACCAACAGTCAGTGGGCCAGAGAGGCTACCTGCTTG 3600
QY 1201 ProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAlaIleGlu 1220
DB |||||
DB 3601 CCAGAGAGCAGCAGTGGCTTCTGTCTCTGCTCAGACCTCCAGCCAGCAGCTGCTATAGG 3660
QY 1221 GluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSer 1240
DB |||||
DB 3661 GAGGCCCTTGGGAAAGCTGCCAGCCCAAGCCCTGTGTCTCTTAACAGTCAGGCCCAAGC 3720
QY 1241 LeuProArgAlaProValSerSerAlaGlnValPro 1252
DB |||||
DB 3721 CTTCCAGGGGCCCTGTCTCTCTCTGCTCAGGTCCCC 3756
```

RESULT 2

AAI67204

XX AAI67204 standard; DNA; 3753 BP.

AC AAI67204;

XX 11-FEB-2002 (first entry)

XX Nucleotide sequence of GSK gene Id 27142.

XX Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;
XX antidiarrheal; nontropic; neuroprotectant; hypotensive; hypertensive;
XX cytosolic; cerebroprotective; vasotropic; human; ds.

XX Homo sapiens.

XX WO200172961-A2.

XX 04-OCT-2001.

XX 22-MAR-2001; 2001WO-US09226.

XX 24-MAR-2000; 2000US-192158P.

XX 28-MAR-2000; 2000US-192668P.

XX 27-APR-2000; 2000US-200166P.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX.

XX Agarwal P, Murdoch PR, Rizvi SK, Smith RP, Xiang Z, Kabnick KS;

XX Lai Y;

XX MPI; 2001-639223/73.

XX P-PSDB; AAG65914.

XX Isolated polypeptides, which may be peptide hormones, which are
XX identified by high throughput genome-based biology which identifies
XX genes and gene products as therapeutic targets for treatment of
XX diseases such as diabetes and cancer -

XX Claim 2; Page 55-56; 99pp; English.

XX The invention provides polypeptides (AAG65886-65918) which may be peptide
XX hormones (including insulin, growth hormones, chemokines, cytokines,
XX neuropeptides, integrins, kallikreins, laminins, melanins, natruiretic
XX hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,
XX secretogranins, selectins, thromboglobulins, thymosins) identified by
XX high throughput genome-based biology and polynucleotides (AAI67176-67208)
XX encoding them. The polypeptides can be expressed by standard recombinant
XX methodology. The polypeptides are useful in the treatment of disease such
XX as diabetes, breast-, prostate-, colon cancer and other malignant tumors,
XX hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,
XX asthma, manic depression, dementia, delirium, mental retardation,
XX Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental
XX or sexual development disorders, and dysfunctions of the blood cascade
XX system including those leading to stroke. The polynucleotides may be used
XX as diagnostic reagents through detecting mutations in the associated gene
XX and for chromosome localization and for tissue expression studies. The
XX polypeptides and polynucleotides may also be used as vaccines.

XX Sequence 3753 BP; 714 A; 1220 C; 1152 G; 667 T; 0 other;

Alignment Scores:

Pred. No.:	1,06e-266	Length:	3753
Score:	5746.00	Matches:	1098
Percent Similarity:	92.09%	Conservative:	54
Best Local Similarity:	87.77%	Mismatches:	96
Query Match:	87.01%	Indels:	5
DB:	22	Gaps:	2

US-09-754-997A-2 (1-1252) x AAI67204 (1-3753)

QY 1 MetAlaArgAlaAspThr-----GlyArgGlyLeuLeuValLeuThrPheCysLeuLeu 18

Qy	739	ThrProAspLeuProIleGlnArgGlyProProLeuProAlaHisValHisAlaGlu	758
Db	2218	GCACCAAGCATGCTATTCAGAGGGACCAACCCTGCCTCCAGCCACGCTCCATGCGAA	2277
Qy	759	SerAsnSerSerThrSerIleTrrLeuArgTrrIlyValyBProAspPheThrThrValIys	778
Db	2278	TCAAAAGAGCTCCACATCCATCTGGCTTCGGTGGAAAGCCAGATTTCACCAACAGTCAAG	2337
Qy	779	IleValAsnTyrThrValArgPheGlyProTrrIglYLeuArgAsnAlaSerLeuValThr	798
Db	2338	ATTGTCACTACTGTGGCTTCAGCCCCCTGGGGCTCAGGAATGCCTCCCTGGTCACC	2397
Qy	799	TyrTyrThrSerSerGlyGluAspIleLeuIleGlyLeuIysProPheThrIysTyr	818
Db	2398	TATTACACCAAGTTCCTGGAGAAGACATCCTCATTTGGGGCTTGAAGGCCATTTCACCAAAATAC	2457
Qy	819	GluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGlu	838
Db	2458	GAGTTTGCAGTCAGATCTCAAGCGGTGGACATGATGGGCTTTTCGGCTCTGTGGTGGAG	2517
Qy	839	ArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeu	858
Db	2518	CGCTCCACCCCTGCCTGACGGGCCCTCACACCCCCATCCGACCTGCGACTGAGCCCCCTG	2577
Qy	859	ThrProSerThrValArgLeuHisTrrCyProProThrGluProAsnGlyGluIleVal	878
Db	2578	ACACCGTCCACGGTTCGGCTGCATGTGTGCCCCCCCACAGAGGCCCAACGGGGAGATCGTG	2637
Qy	879	GluTrrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrrThrLeuLeu	898
Db	2638	GAGTATCTGATCTCTGTACAGCAACCAACAGCAGCTTGAGCACCACTGTGCTC	2697
Qy	899	ThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyr	918
Db	2698	ACCACGACGGGAACATCTTCAGTGTCTGAGTCCATGGCTTGAGAGCGACACTCGGTAC	2757
Qy	919	PhePheIysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAsp	938
Db	2758	TTCTTCAAGATGGGGGCGCACAGAGGTGGGACCTGGGGCTTTTCTCCGCCCTGCAGGAT	2817
Qy	939	ValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIle	958
Db	2818	GTGATCACGCTCCAGGAGAGACTGTCAAGCTCGCTGGACATGCACTCAGTCAGTCGGGATC	2877
Qy	959	IleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArg	978
Db	2878	ATCGTGGGTGTCTGCCCTGGGCTCTCTGCTCTCTGCTGCGATGTGTGCTGGCTCGC	2937
Qy	979	GlnSerSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnPro	998
Db	2938	CGCAGCCCCCACAGGGAATCCCTCCAGGCTGTCTCCACCGGCCACCCCGGGAAATCCC	2997
Qy	999	AlaLeuTyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGlu	1018
Db	2998	CGCGTGTCTCCAGAGCTGGCTTGGCCCCCCCCAGCCCCCATGAATTGGAG	3057
Qy	1019	SerLeuValHisProArgProGlnAspTrrSerProProProSerAspValGluAspIys	1038
Db	3058	TCCCTTGTGACCCCCCATCCACAGACTGGTTCGCCCGCACCTCAGACGCTGGAGCAAGG	3117
Qy	1039	AlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArg	1058
Db	3118	GCTGAAGTGACACGCTTATGGGTGGCGGTGTTCCTGAAGGCCCGAGTCACTCCAAAGA	3177
Qy	1059	LysIleSerTrrAlaGlnAlaGlyGlyProAsnTrrAlaGlySerTrrAlaGlyCysGlu	1078
Db	3178	AAGATCTCTGGGCTCAACCAAGCGGGCTGAGCTGGCTGGTCTCTGGCAGGCTGTGAG	3237
Qy	1079	LeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGly	1098
Db	3238	CTGCCCCCAG---GAGCGCCCCGCGCTGTGACCGGGGCCCTGTGCTGCCCTGTGCGA	3294

Qy	1099	ThrGlyGlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleIysSerAsnGlyArg	1111
Db	3295	ACTGGGCAGACGCTGTGTCTGCAGGCTCTGGTGTACGACGCCAATAAAGGGCAATGGGAGG	3354
Qy	1119	LYsLYeProSerProAlaCYsArgAanGlnValGluAlaGluValIleValHisSerAsp	1138
Db	3355	AAGAAGTCACCCCGACGCTGCAGGAACCGAGTGGAGGCTGAGTCAATTGTCCACTCTGAC	3414
Qy	1139	PheGlyAlaSerLYsGlyCYsProAspLeuHisLeuGlnAspLeuGluProGluGluPro	1158
Db	3415	TTTAGTGCACTTAACGGGAACCTCGACCTCCATCTCCAAGACCTGGAGCCTGAGGACCCC	3474
Qy	1159	LeuThrAlaGluThrLeuProSerThrSerGlyAlaValaAspLeuSerGlnGlyAlaAsp	1178
Db	3475	CTCGCTCCAGAGGCTCTCGATCTCATCTCGGGTGTGGGGATCCAGGGCAGGGGGCAGCC	3534
Qy	1179	TrpLeuGlyArgGluLeuGlyGlyCYsGlnProThrThrSerGlyProGluArgLeuThr	1198
Db	3535	TGGCTGACACGGAGTGTGGAGAGGTGTGACTGGCAGCCCCCGGCCAGACACTTACC	3594
Qy	1199	CysLeuProGluAlaAlaSerAlaSerCYsSerCYsSerAspLeuGlnProSerThrAla	1218
Db	3595	TGCTTGCCAGAGCAGCAGTGTCTTCTGCTCTACCCGGACCTCCAGCCAGCGAGGTG	3654
Qy	1219	IleGluGluAlaProGlyLYsSerCYsGlnProLYsAlaLeuCYsProLeuThrValSer	1238
Db	3655	CTAGAGGAGACCCCTCTGGAGATAGCTGCAGCTCAAAATCCCCCTGCCCTCTAGGAGCAGC	3714
Qy	1239	ProSerLeuProArgAlaProValSerSerAla	1249
Db	3715	CCAGGCTGCCAGATCCCCGGTCTCTCTCTCT	3747
RESULT 3			
ABK11101	ID ABK11101 standard; DNA; 3741 BP.		
XX AC	ABK11101;		
XX DT	05-JUN-2002 (first entry)		
XX DE	DNA encoding human NOVI protein, homologue of NOPE/PUNC Ig proteins.		
XX KW	Human; NOVI-associated disorder; developmental disorder; blood disorder;		
XX KW	endocrine disorder; vascular disease; gastrointestinal disorder; cancer;		
XX KW	respiratory disorder; inflammatory disorder; reproductive disorder;		
XX KW	neurodegenerative disorder; autoimmune disorder; infectious disease;		
XX KW	cardiovascular disorder; cell signal processing; PUNC immunoglobulin;		
XX KW	metabolic pathway modulation; NOPE; Ig; NOVI; gene; ds.		
XX OS	Homo sapiens.		
XX FH	Key	Location/Qualifiers	
XX CDS	1..3741	/*tag= a	
XX FT	/partial	/product= "NOVI"	
XX FT	/note= "This sequence lacks a stop codon"		
XX FN	WO200206329-A2.		
XX PD	24-JAN-2002.		
XX PF	18-JUL-2001; 2001WO-US22709.		
XX PR	18-JUL-2000; 2000US-218870P.		
XX PR	18-JUL-2000; 2000US-218875P.		
XX PR	18-JUL-2000; 2000US-218901P.		
XX PR	24-JUL-2000; 2000US-220273P.		
XX PR	26-JUL-2000; 2000US-220912P.		
XX PR	27-JUL-2000; 2000US-221233P.		
XX PR	28-JUL-2000; 2000US-221650P.		
XX PA	(CURA-) CURAGEN CORP.		

QY 518 LeuValHisThrLeuAspValProSerAlaAlaProGlnLeuThrLeuSerSerPro 537
DB 1555 CTGGTGCAACACTGGATGATCTCCAGAGCAACACACGAGCCTGAGCACAGTGGACCTTG 2691
QY 538 AsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGln 557
DB 1615 AACCCCTTCGGAATCAGGGTGGCGTGGCTGCCCTTGCCTCCAGCTGAGCAATGGGCG 1674
QY 558 ValLeuLysTrpLysIleGluTrpGlyLeuGlyLysGluAspGlnValPheSerThrGlu 577
DB 1675 GTGGTGAAGTACAGATAGAAACGGTTTGGGAAGGAGATCAGATTTCCTACTGAG 1734
QY 578 ValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTrpArg 597
DB 1735 GTGCGAGGAAATCAGACACAGCTTATGCTGAACCTCGCTTTCAGCCAAACAGAGTGTATCGA 1794
QY 598 ValArgIleSerAlaGlyThrGlyAlaGlyTrpGlyValProSerGlnTrpMetGlnHis 617
DB 1795 GTACGGATTTCGGCTGGTACAGAGCCGGCTTTCGGGGCCCTCCAGTGGATGCATCAC 1854
QY 618 ArgThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysVal 637
DB 1855 AGGAGCCCCAGTATGCACACACAGGCCATGTCCTTTTGGCTTGCAGAGTTGAAGGTG 1914
QY 638 ArgAlaLysMetGluSerLeuValValSerTrpGlnProProProHisProThrGlnIle 657
DB 1915 CAGGCAAAAGATGGAGTCCCTGGTGTGTATGGCAGCCACCCCTCACCCACCCAGATC 1974
QY 658 SerGlyTrpLysLeuTrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArg 677
DB 1975 TCTGGCTACAAACTATATTGGCGGGAGGTGGGGCTGAGGAGGAGGCCAATGGCGATCGC 2034
QY 678 ProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysLys 697
DB 2035 CTGCCAGGGGGCGTGGAGACACAGGCTTGGGATGTGGGGCTGTCCGGCTCAAGAGAA 2094
QY 698 ValLysGlnTrpGluLeuThrGlnLeuValProGlyArgProTrpGlyValLysLeuVal 717
DB 2095 GTGAAGCAGTATGAGCTGACCCAGCTAGTCCCTGGCGGCTGTATCGAGGTGAAGCTGGT 2154
QY 718 AlaPheAsnLysHisGluAspGlyTrpAlaAlaValTrpLysGlyThrGluLysAla 737
DB 2155 GCTTTTCACAAACATGAGATGGCTATGACAGAGTGTGGAGGGCAAGACGAGAGAGCGC 2214
QY 738 ProThrProAspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAla 757
DB 2215 CCGGCACACAGATGCTTATCCAGAGGGGACACCCCTGCTCCAGCCACGCTCCATGGC 2274
QY 758 GluSerAsnSerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrVal 777
DB 2275 GAATCAAAACAGCTCCACATCCATCTGGCTTCGGTGGAAAAAGCCAGATTTCACACAGTC 2334
QY 778 LysIleValAsnTrpThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuVal 797
DB 2335 AAGATTGTCAACTACATCTGTGGCTTTCAGCCCTTGGGGGCTCAGGAAATGCTCCCTGGTGC 2394
QY 798 ThrTrpTrpThrSerSerGlyLysAspIleLeuIleGlyLysLeuLysPheThrLys 817
DB 2395 ACCATATAC---AGTTCTGAGAAAGACATCCTCATTTGGCGGCTTGAAGCCATTCACAAA 2451
QY 818 TyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValVal 837
DB 2452 TAGCAGTTTTCAGTGCAGTCTACGCGGTGGACATGATGGGCTTTCGGCTCTGTGGTG 2511
QY 838 GluArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerPro 857
DB 2512 GAGCGCTCCACCCCTGCTGACCGCTCCACACCCCTCCAGACCTCCGAGCTGAGCCCTC 2571
QY 858 LeuThrProSerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIle 877
DB 2572 CTGACACCGTCCACCGTGTGGCTGCACTGGTGGCTGGTGGCTGGTGGCTGGTGGCTGGT 2631
QY 878 ValGluTrpLysLeuIleLeuTrpSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeu 897

DB 2632 GTGGAGTATCTGATCTCTGTATCAGAGCAACACACGAGCCTGAGCACAGTGGACCTTG 2691
QY 898 LeuThrThr-----GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAsp 915
DB 2692 CTCACACGAGGGTGGAGGAAACATCTTCACTGCTGAGGTCCATGGCTGAGAGGCCAC 2751
QY 916 ThrArgTrpPhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArg 935
DB 2752 ACTCGTACTTCTTCAAGATGGCGCGGCACAGAGGTGGGACCTGGGCTTCTTCCGCGC 2811
QY 936 LeuGlnAspValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaVal 955
DB 2812 CTGAGGATGTGATCAGCTCCAGGAGAAGCTGTGAGATCGCTGGACATCAGCTCAGTC 2871
QY 956 ThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAla 975
DB 2872 ACGGGCATCATCTGGGTGTCTGCTGGGCTCTCTGCTCTCTGGCTTGCATGTGTGCT 2931
QY 976 GlyLeuArgGlnSerSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrPro 995
DB 2932 GGCCTGGCGCGCAGCCCCACAGGGAATCCCTCCAGGCTGTCTCCACCGCCACCCCTC 2991
QY 996 GlyAsnProAlaLeuTrpThrArgAlaArgLeuGlyProProSerValProAlaAlaHis 1015
DB 2992 GGGAAATCCCGGCTGTACTCCAGAGCTCGGCTTGGCCCCCAGCCCGCCAGCTGCCAT 3051
QY 1016 GluLeuGluSerLeuValHisProArgProGlnAspTrpSerProProSerAspVal 1035
DB 3052 GAATTGGAGTCCCTTGTGCACCCCATCCAGGACTGTGTCCTCCGCCACCTTCAGACGTG 3111
QY 1036 GluAspLysAlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHis 1055
DB 3112 GAGGACAGGGCTGAAGTGCACAGCTTATGGTGGCGGTCTTCTGAAGGCGGAGTCCAC 3171
QY 1056 SerLysArgLysIleSerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAla 1075
DB 3172 TCCAAAAGAAAGGTAAAGT---GCTCAACAAAGCGGCTGAGCTGGGCTGTGTTCTTGGGCA 3228
QY 1076 GlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuPro 1095
DB 3229 GGTGTGAGTGGCCCGCAG---GAGGCCCCCGCGCGGTCTGACCCGGGCTCTGCTGCC 3285
QY 1096 ProAlaGlyThrGlyGlnThrLeuLeuLeuGlnAlaLeuValTyrAspGlyIleLysSer 1115
DB 3286 CCTGTGTAAGTGGGACAGCGCTGTGTGTCAGGTTCTCTGCTCTGAT-----CAGGCGC 3339
QY 1116 AsnGlyArgLysLysProSerProAlaCysArgAsnGlnValGluAlaGluValIleVal 1135
DB 3340 AATGGAGAGAAAGATCACCCCGCTGCAGGAACACAGGTGGAGGCTGAAGTCAATTGTC 3399
QY 1136 HisSerAspPheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluPro 1155
DB 3400 CACTCTGACTTTAGTGCATCTAACCGGAACCTGACCTCCATCTCCAGACCTGAGGCT 3459
QY 1156 GluGluProLeuThrAlaGluThrLeuProSerThrSerGlyValAlaValAspLeuSerGln 1175
DB 3460 GAGGACCCCTGCTCCAGAGGCTCTCATCTCATCTCGGGTGTGGGATCCAGGGCAG 3519
QY 1176 GlyAlaAspTrpLeuGlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyProGlu 1195
DB 3520 GGGGCGAGCTGCTGGACAGAGGAGTGGAGGGGTGTGAGCTGGCAGCCCGCGGCGCAGAC 3579
QY 1196 ArgLeuThrCysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnPro 1215
DB 3580 AGACTTACTGCTTCCAGAGGAGGAGGAGGAGTGTCTCTGCTCTCTACCCGAGCTCCAGCCA 3639
QY 1216 SerThrAlaIleGluAlaProGlyLysSerGlnProLysAlaLeuCysProLeu 1235
DB 3640 GCGAGGTGCTAGAGGAGGAGGAGGAGGAGTGTCTCTGCTCTCTACCCGAGCTCCAGCCA 3699
QY 1236 ThrValSerProSerLeuProArgAlaProValSerSerAla 1249

Db 3700 GGAGCCAGCCAGCGCTGCCAGATCCCGGTCTCTCTCTCTCT 3741

RESULT 4

AAD10022

ID AAD10022 standard; cDNA; 2796 BP.

XX AAD10022;

XX 12-SEP-2001 (first entry)

XX Mouse Nope (neighbour of punc ell) extracellular domain cDNA.

XX Mouse; Nope; neighbourhood of punc ell; cytosolic; neuroprotective; vaccine;
 XX gene therapy; cerebroprotective; colonic cancer; mental retardation;
 XX tumour suppressor; chromosome 9; transgenic animal; genetic disorder;
 XX obesity; Bardet-Biedl syndrome; autosomal recessive disorder; retinitis;
 XX polydactyly; pigmentosa; hypogonadism; hypertension; diabetes mellitus;
 XX renal anomaly; cardiovascular anomaly; extracellular domain; ss.

XX Mus musculus.

XX Key Location/Qualifiers
 XX 1..2796

XX /*tag= a

XX /product= "Mouse Nope extracellular domain"

XX /note= "CDS does not include start and stop codon"

XX /partial

XX WO200149714-A2.

XX 12-JUL-2001.

XX 26-OCT-2000; 2000WO-US29698.

XX 04-JAN-2000; 2000US-0174496.

XX 19-MAY-2000; 2000US-0205789.

XX (NEUR-) NEUROSCIENCES RES FOUND INC.

XX Salbaum JM;

XX WPI; 2001-441846/47.

XX P-PSDB; AAE05252.

XX Murine Nope polypeptides and nucleic acids useful for preventing,

XX diagnosing and treating colonic cancer and Bardet-Biedl syndrome -

XX Claim 11; Page 77-81; 99pp; English.

XX The present invention relates to Nope (neighbour of punc ell) which is
 CC used in the prevention, treatment and diagnosis of diseases associated
 CC with inappropriate Nope expression such as cancers especially colonic
 CC cancer and genetic disorders, as Nope is thought to be a tumour
 CC suppressor. Nope gene is located on chromosome 9 and is used in gene
 CC therapy. Nope is used as vaccine. Nope gene may be administered to treat
 CC diseases by rectifying mutations or deletions in a patient's genome that
 CC affect the activity of Nope by expressing inactive proteins or to
 CC supplement the patients own production of Nope polypeptides. Nope gene
 CC is used to study the expression and function of Nope polypeptides and
 CC their role in metabolism through the creation of transgenic animal
 CC models. The anti-Nope antibodies and Nope antagonists may also be used
 CC to down regulate Nope expression and activity for the treatment of
 CC Bardet-Biedl syndrome which is an autosomal recessive disorder
 CC characterised by mental retardation, obesity, polydactyly, retinitis
 CC pigmentosa and hypogonadism. Patients with Bardet-Biedl syndrome have a
 CC high incidence of hypertension, diabetes mellitus and renal and
 CC cardiovascular anomalies. The present sequence is mouse Nope (neighbour
 CC of punc ell) extracellular domain cDNA.

XX SQ Sequence 2796 BP; 591 A; 849 C; 821 G; 535 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 2.1e-226 Length: 2796

Score: 4903.00 Matches: 932

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 74.24% Indels: 0

DB: 22 Gaps: 0

US-09-754-997A-2 (1-1252) x AAD10022 (1-2796)

QY 22 GlyGluLeuProLeuProGlnGluThrValLysLeuSerCysAspGluGlyProLeu 41

DB 1 GGAGAGCTGCCATTGGCCCGCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGACCCCTG 60

QY 42 GlnValIleGluGlyProGlnGlnAlaValValLeuAspCysThrLeuGlyAlaThrAla 61

DB 61 CAAGTGATCTGGCGCCCTGAGCAGGCTGTGGTGTGACTGCACATTTGGGGGGCTACAGCT 120

QY 62 AlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGluHisGlu 81

DB 121 GCTGGCCCTCCGACCGGTGACATGGACAGGATGAGATGGAGACACTGTACTAGAGCATGAG 180

QY 82 AsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAsp 101

DB 181 AACCTGCACCTGCTACCCCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAGAGAGAC 240

QY 102 SerAspAspGluAlaLeuArgIleTrpLysValThrGluGlySerTyrSerCysLeu 121

DB 241 AGCGATGATGAGAGAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGTCTG 300

QY 122 AlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThrLeu 141

DB 301 GCCCACAGCCCGCTAGGATGGTGGCCAGCCAGGTGTGCTGTGGTCAAGCTTGGCCACATC 360

QY 142 GluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArgPhe 161

DB 361 GAAGACTTCTCTGACCCCGAGTCCAGATGTGGAGGAGNACGGGACAGACACGCTTT 420

QY 162 GluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnValThr 181

DB 421 GAATGCCACACCAAGGGCTTCCAGCCCCCATCATTTCTTCCAGATGGCTCTCTCAGATGTC 480

QY 182 ValProGluGluProArgIleThrLeuProLysTrpLeuGlnIleLeuAspVal 201

DB 481 GTGCCCTGAGAGAGCCCGGCTCATCTCTTCCAGATGGCTCTCTCAGATGTC 540

QY 202 GlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPhe 221

DB 541 CAGGACAGTGATGACAGGCTCTTACCGCTGCTGGCCACCAATTCAGCCCGCCCAACCATTC 600

QY 222 SerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGln 241

DB 601 AGCCAGGAGCGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAAGGGGCGAG 660

QY 242 AspValValIleValAlaAlaProGluAsnThrValValSerGlyGlnAsnValVal 261

DB 661 GATGTGGTCAATGTGGCAGCCCGCAGAGNACCAACCGTAGTGTCTGGCAGAGATGTAGTG 720

QY 262 MetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAspGly 281

DB 721 ATGGAGTGTGGCCCTCTGCTGACCCCAACCCCTTTTGTCTTGGTCCGACAGGATGGA 780

QY 282 LysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAla 301

DB 781 AAGCCTATCTCCAGGATGTGCTGTCTGGCCCGGACCAATCTACTCATCGCCAGCGCG 840

QY 302 GlnProArgHisSerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPhe 321

DB 841 CAGCCTCGGCACTCTGGAGTGTATGTCTGCGGAGCAACAAGGCCCTTCACGCGTGCATTC 900

QY 322 AlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGlu 341

DB 901 GCCACTGCGGCTGTGAGCTCCGAGTGTGTGCTGCCCCAGCCCATCTCGGAGGACCCGAG 960

QY 342 AlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluPro 361

PD 12-SEP-2002.
 XX
 PF 05-MAR-2002; 2002WO-US05095.
 XX
 PR 05-MAR-2001; 2001US-0799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich RW, Auendi V, Zhang J, Zhao QA, Ren F;
 PI Yue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Dmanac RT;
 XX
 XX WPI; 2002-759812/82.
 DR P-PSDB; ABP69002.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
 PT platelet or coagulation disorders -
 XX
 PS Claim 1; SEQ ID NO 101; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences
 CC (ABZ1119-ABZ12066) or their mature protein coding portion, active domain
 CC coding protein or complementary sequences. The polynucleotides are useful
 CC for identifying expressed genes or for physical mapping of human genome.
 CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
 CC weight markers, as a food supplement, for generating antibodies, in
 CC medical imaging, screening and diagnostic assays and for treating
 CC cell-proliferative disorders (cancer), neurodegenerative diseases
 CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
 CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
 CC disorders, platelet or coagulation disorders, wound, burns, incision,
 CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
 CC parasitic), arthritis, etc.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3450 BP; 681 A; 1050 C; 1067 G; 652 T; 0 other;
 XX

Alignment Scores:
 Pred. No.: 2,46e-200 Length: 3450
 Score: 4361.50 Matches: 822
 Percent Similarity: 92.55% Conservative: 35
 Best Local Similarity: 88.77% Mismatches: 68
 Query Match: 66.04% Indels: 1
 DB: 24 Gaps: 1

US-09-754-997A-2 (1-1252) x ABZ11219 (1-3450)

QY 324 AlaAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeu 343
 DB 163 TCAGCGCTGCCCTCGCGCCCTAGCGGCTCCCGCCATCATCAGCGCGCCGAGGCGCTG 222
 QY 344 SerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgPro 363
 DB 223 TCSCGAGCGCGGAGCAGCAGCGCGCTTCGTGTGCGCGCTGCGGAGCGCGGCCA 282
 QY 364 AlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgVallysValGln 383
 DB 283 CGCGTGGCTGGCTGCACACAGCGGCGCGCTCGCGGCCAACCGCGCGCTCAAGGTCCAG 342
 QY 384 GlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGln 403
 DB 343 GCGCGCGGTGGCAGCTGGTTCATCACACATCGGCTGCAGGACGCGCGCTACTACCAG 402
 QY 404 CysValAlaGluHisSerAlaGlyThrAlaCysAlaAlaAlaProLeuAlaValVal 423
 DB 403 TGGTGGCTGAGAACAGCGCGGGAATGCGCGTGGCTGCGCGCTGCGCGCGTGGTGGTG 462
 QY 424 ArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSer 443

DB 463 CCGCAGGGGCTGCCCGCGCGCCCGCGGGTCACTGCTACGCCACTGAGCAGCTCGCT 522
 QY 444 ValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeu 463
 DB 523 GTGTGGTGGCTGGGAGCGCGCCGAGATGCACAGCGAGCAGATCATCGCTTCTCTCTC 582
 QY 464 HisTyrGlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAsp 483
 DB 583 CACTACCAAGAGCGCGGCGCATGGACATGTGGAAATCCAGTTTGCAGTGAACACGAC 642
 QY 484 ThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValVal 503
 DB 643 ACCCAGAACTACAGGTCGGGACCTGGAAACCCCAACACAGATATAGAGTTCTACGTGG 702
 QY 504 AlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAsp 523
 DB 703 GCTTACTCCAGCTGGAGCGCAGCGCACTCCACCCAGCACTGGTGCACACACTGGAT 762
 QY 524 AspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArg 543
 DB 763 GATGTCCCGAGTGCAGCACCCCGAGCTCTCCCTGTCAGCCCAACCCCTTCGGACATCAG 822
 QY 544 ValAlaTrpLeuProLeuProSerSerSerSerSerSerSerSerSerSerSerSer 563
 DB 823 GTGGCGTGGCTGCTCCCTGCCCCCGCCAGCAATGGCGAGTGGTGAAGTACAGATA 882
 QY 564 GluTyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThr 583
 DB 883 GAATACGGTTTGGGAAAGGAGATCAGATTTCTTCTAGAGTGGGAGGAAATGAGACA 942
 QY 584 GlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGly 603
 DB 943 CAGCTTATGCTGAATCGCTTCAGCCAAACAAAGGTATCGAGTACGGATTTCCGCTGGT 1002
 QY 604 ThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHis 623
 DB 1003 ACAGCAGCGCGCTCGGGGCGCCCTCCCAAGTGGATGCATCAGAGGACGCCATATGCA 1062
 QY 624 AsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSer 643
 DB 1063 AACCCAGGCCATGTCTCTTTCCCTGCGAGATTGAGGTGCGAGCAAGATGGAGTCC 1122
 QY 644 LeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTyrIleLeuTyr 663
 DB 1123 CTGGTGTGTCTGTCAGCGCACCCCTCACCACCCAGCATCTCTGGCTACAAATATAT 1182
 QY 664 TrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProProGlyGlyArgGly 683
 DB 1183 TGGCGGGAGGTGGGGCTGAGGAGGAGGCCAATGGCGATCGCTGCGAGGGGCGCGTGA 1242
 QY 684 AspGlnAlaTrpAspValGlyProValArgLeuLysValLysValLysGlnTyrGluLeu 703
 DB 1243 GACCCAGGCTTGGAGTGGGGGCTGTCCCGCTCAGAGAGTGAAGTGAAGCATATAGCTG 1302
 QY 704 ThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGlu 723
 DB 1303 ACCCAGCTAGTCCCTGGCGCGCTGTACGAGGTGAAGTCTGTGGCTTTCACAAACATGAG 1362
 QY 724 AspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProLeuPro 743
 DB 1363 GATGGCTATGCAGCATGTGGAAAGGCGCAAGAGAGCGCGCGCGCACGACATGCT 1422
 QY 744 IleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerSerSerSer 763
 DB 1423 ATCCAGAGGGGACCCCTGCTCCAGCCCGCTCAGCGGAAATCAACAGCTCCACA 1482
 QY 764 SerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsnTyrThr 783
 DB 1483 TCCATCTGCTTGGTGGGAAAGGAGATTTCCACCAAGTCAAGATTTCACTACTACT 1542
 QY 784 ValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyrThrSerSer 803

Db 1543 GTGCGCTTTCAGCCCTGGGGGCTCAGGAATGCTCCCTGGTCACTATTACACCAAGTTCT 1602
Qy 804 GlyGluAspIleLeuIleGlyGlyLeuValProPheThrLeuValGluPheAlaValGln 823
Db 1603 GGAGAAGACATCCTCATTTGGCGGCTTGAAGCAATTCACAAATAGAGTTTGAGTGGAG 1662
Qy 824 SerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThrLeuPro 843
Db 1663 TCTCAGCGGTGGACATGATGGGCTTTTGGCTCTGTGGTGGAGGGCTCCACCCCTGGCT 1722
Qy 844 AspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSerThrVal 863
Db 1723 GACCGGCTTCACACCCCTCATCCGACCTGGAGCTGAGCCCTGACACCGCTCCACGGTT 1782
Qy 864 ArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyrLeuIleLeu 883
Db 1783 CGGCTGCACTGGTGGCCCCCAGAGCCCAAGGGAGATCGTGAGTATCTGATCTGT 1842
Qy 884 TyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThrGluGlyAsn 903
Db 1843 TACAGCAGCAACACACAGCAGCTGAGCACACAGTGGACCTTCTCACCAGCGAGGAAAC 1902
Qy 904 IlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLeuMetGly 923
Db 1903 ATCTTCAGTGTGAGTCCATGGCTGGAGAGCGACACTCGGTACTTCTTCAAGATGGG 1962
Qy 924 AlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThrLeuGln 943
Db 1963 GCGGCGACAGAGTGGGACCTGGGCTTTCTCCCGCTGCAGGATGTGATCAGCTCCAG 2022
Qy 944 GluThrPheSerAspSerLeuAspValHisAlaValThrGlyIleLeuValGlyValCys 963
Db 2023 GAGAAGCTCTCAGACTCGCTGGACATGCACTCAGTCACGGGATCATCTGGTGGTCTGC 2082
Qy 964 LeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSerSerHisArg 983
Db 2083 CTGGGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2142
Qy 984 GluAlaLeuProGlyLeuSerSerGlyThrProGlyAsnProAlaLeuThrThrArg 1003
Db 2143 GAATCTCTCCAGGCTGTCTCCACCCCGCCGGAATCCCGGCTGTACTCCAGA 2202
Qy 1004 AlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeuValHisPro 1023
Db 2203 GCTGGCTTGGCCCCCAGCCCCCAGCTGCCCCATGATGAGTCCCTTGTGACCC 2262
Qy 1024 ArgProGlnAspTrpSerProProSerAspValGluAspLeuAlaGluValHisSer 1043
Db 2263 CATCCCCAGCACTGGTCCCGCCACCCCTCAGACGTCGAGGACAGGCTGAAGTCACAGC 2322
Qy 1044 LeuMetGlyGlySerValSerAspCysArgGlyHisSerLeuArgGlySerThrAla 1063
Db 2323 CTATGGGTGGGGGTGTTCTGAAGCCCGGAGTCACTCCAAAGAAAGATCTCCTGGGCT 2382
Qy 1064 GlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySer 1083
Db 2383 CAACCAAGCGGCTGAGCTGGGCTGTTCTTGGCAGGCTGTGAGTGGCCCCAG --GCA 2439
Qy 1084 GlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeu 1103
Db 2440 GGGCCCCGGCGCTCTGACCCCGGCGCTGCTGCCCTGCTGGAACCTGGGAGAGCGCTG 2499
Qy 1104 LeuLeuGlnAlaLeuValTrpAspGlyIleLeuSerAsnGlyArgGlyLeuProSerPro 1123
Db 2500 TTGCTGACGGCTCTGTGTGACGACGCAATAAGGGCAATGGGAGGAAGATCACCCTCA 2559
Qy 1124 AlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGlyAlaSerLeu 1143
Db 2560 GCTGAGAGAACAGGTGAGGCTGAGTCATGTGTCTGCTGCTGCTGCTGCTGCTGCTG 2619
Qy 1144 GlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThrAlaGluThr 1163
Db 2620 GGGACCCCTGACCTCATCTCCAAAGACCTGGAGCTGGAGCCCTGAGGACCCCTGCTCCAGAGGCT 2679

RESULT 6

AAI67203

ID AAI67203 standard; DNA; 2223 BP.

XX

AC AAI67203;

XX

DT 11-FEB-2002 (first entry)

XX

DE Nucleotide sequence of GSK gene id 27142.

XX

KW Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;
KW antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;
KW cytosolic; cerebroprotective; vasotropic; human; ds.

XX

OS Homo sapiens.

XX

FN WO200172961-A2.

XX

PD 04-OCT-2001.

XX

PF 22-MAR-2001; 2001WO-US09226.

XX

PR 24-MAR-2000; 2000US-192158P.

PR

PR 28-MAR-2000; 2000US-192668P.

PR

PR 27-APR-2000; 2000US-200166P.

XX

(SMK) SMITHKLINE BEECHAM CORP.

PA

(SMK) SMITHKLINE BEECHAM PLC.

XX

PI Agarwal P, Murdoch PR, Rizvi SK, Smith RP, Xiang Z, Kabnick KS;

PI

Lai Y;

XX

WPI; 2001-639223/73.

DR

P-PSDB; AAG65913.

XX

Isolated polypeptides, which may be peptide hormones, which are

PT

identified by high throughput genome-based biology which identifies

PT

genes and gene products as therapeutic targets for treatment of

PT

diseases such as diabetes and cancer

XX

Claim 2; Page 54-55; 99pp; English.

PS

XX

CC

CC

CC

CC

CC

CC

CC

The invention provides polypeptides (AAG65886-65918) which may be peptide hormones (including insulin, growth hormones, chemokines, cytokines, neuropeptides, integrins, kallikreins, lamins, melanins, natriuretic hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins, secretogranins, selectins, thromboglobulins, thymosins) identified by high throughput genome-based biology and polynucleotides (AAI67176-67208) encoding them. The polypeptides can be expressed by standard recombinant methodology. The polypeptides are useful in the treatment of disease such as diabetes, breast-, prostate-, colon cancer and other malignant tumors, hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,

CC asthma, manic depression, dementia, delirium, mental retardation,
 CC Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental
 CC or sexual development disorders, and dysfunctions of the blood cascade
 CC system including those leading to stroke. The polynucleotides may be used
 CC as diagnostic reagents through detecting mutations in the associated gene
 CC and for chromosome localization and for tissue expression studies. The
 CC polypeptides and polynucleotides may also be used as vaccines.

XX SQ Sequence 2223 BP; 440 A; 660 C; 735 G; 388 T; 0 other;

Alignment Scores:

Pred. No.: 2,84e-143 Length: 2223
 Score: 3167.50 Matches: 620
 Percent Similarity: 90.33% Conservative: 34
 Best Local Similarity: 85.64% Mismatches: 54
 Query Match: 47.96% Indels: 18
 DB: 22 Gaps: 2

US-09-754-997A-2 (1-1252) x AA167203 (1-2223)

Qy 17 LeuLeuSerAlaArgGlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCys 36
 Db 61 CTCTCTCTATGCAGGGAGCTGCTGTGCCCCAGGAGAGCTGTGGAGCTGAGCTGT 120
 Qy 37 AspGluGlyProLeuGlnValLysLeuGlyProGlnAlaValValLeuAspCysThr 56
 Db 121 GGAGTGGGGCCACTGCAAGTATCTCTGGGCCAGAGCAGCTGCAGTGTAAACTGTAGC 180
 Qy 57 LeuGlyAlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThr 76
 Db 181 CTGGGGCTCTGCTGCGCTGACCCCTGACCCCTGAGGGTGCCTGGAGCAAGGATGGGACACC 240
 Qy 77 ValLeuGluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerPro 96
 Db 241 CTGCTGGAGCAGCAGCACTTACCTCTCTGCCCATGTTCCCTGTGGCTGTCCAGCCCA 300
 Qy 97 LeuGluGlnGluAspSerAspGluAlaLeuArgLysValThrGluGly 116
 Db 301 CTAGCACCACCAATGGCAGTACAGT--CAGTCCCTGAGGCTGTGGGG-GTCATTGAAGGC 357
 Qy 117 SerTyrSerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaVal 136
 Db 358 AACTATTCTGCTGTAGCCCAAGCCCTGAGGCTGTGGAGTGTGGCCAGCAGCTGTCTGTC 417
 Qy 137 LysLeuAlaThrLeuGluAspPheSerLeuHisProGluSerGlnLysValGluAsn 156
 Db 418 AAGCTTCCACACTCGCAGACTTCTCTGACCCCGAGTCTCAGCGTGGAGGAGAAC 477
 Qy 157 GlyThrAlaArgPheGluCysHisThrLysGlyLeuProAlaProLysLeuThrTrpGlu 176
 Db 478 GGGACAGCTGCTTTGAGTGGCCATTTGAAGGGCTGCAGCTCCCATCTACTTGGGAG 537
 Qy 177 LysAspGlnValThrValProGluGluProArgLysLeuThrLeuProLysTrpLeuLeu 196
 Db 538 AAGGACCAAGTGCATTTGCTGAGAGGCTCGGCTCATCTGCTTCCCAACGGGCTCTT 597
 Qy 197 GlnLysLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSer 216
 Db 598 CAGATCTGTGATGTCAGGAGAGTATGCAAGGCCCTTACCCTGCTGGGCCCAACAACCTCA 657
 Qy 217 AlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGlu 236
 Db 658 GCTCGCCAGCACTTACCCAGAGGCGCTTACTAGTGTGGCCACAGAGGCTCTCTGGCG 717
 Qy 237 AlaThrArgGlyGlnAspValValLysLeuAlaAlaProGluAsnThrThrValValSer 256
 Db 718 TCCACCAAGGGGAGGAGGCTGCTTCTGTCAGTGTGGCCACAGAGAACCAACAGTGTCT 777
 Qy 257 GlyGlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrp 276
 Db 778 GGCCAGAGTGTGGTGAATGTGTGGCTCAGCTGACCCCTCAGTGTGGCTGTCTCTCT 837
 Qy 277 ValArgGlnAspGlyLysPheLeuSerThrAspValLysValLeuGlyArgThrAsnLeu 296

Db 838 GTCCGACAGAGCGGAGGCCATCTCCACAGATGTATCTGTCTGGGGCGCACCAACCTA 897
 Qy 297 LeuLeuAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysPro 316
 Db 898 CTAAATTGCCAACCGGAGCGCTGGCACTCGGGGCTCTATGTCTGCGCGCCACACAGCCC 957
 Qy 317 LeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle 336
 Db 958 CGCACGCGGACTTCCGCACTGCGCGCTGAGCTCGGTGTGTGGCGCTCCGCGCATC 1017
 Qy 337 SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg 356
 Db 1018 ACTCAGCGCGCGAGCGCTGTGCGGAGCGGGCGAGCACAGCGGCTGTGTGTGCCGC 1077
 Qy 357 AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyLysLeuArgPro 376
 Db 1078 GCGTCCGGGGAGCGCGGCGCTGCGCTGCTGCTGCAACAACGGGGCGCGCTGCGGCC 1137
 Qy 377 AsnGlyArgValLysValGlnGlyGlySerLeuValLysLeuThrGlnLysLeu 396
 Db 1138 AACGGCGCGTCAAGGTCCAGGGCGCGGTGGCGCTGTGTCTATCACACAGATCGGCTG 1197
 Qy 397 GlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAla 416
 Db 1198 CAGGACCGCGCTACTACCACTGCTGGCTGAGAACAGCGCGGAAATGGCGTGGCTGCC 1257
 Qy 417 AlaProLeuAlaValValValArgGluGlyLeuProSerAlaProThrArgValThrAla 436
 Db 1258 GCGTCTGCTGCGCTGTGTGGTGGCGAGGGCTGCCAGCGCCCCACCGGGTCACTGCT 1317
 Qy 437 ThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlu 456
 Db 1318 AGCCCACTGAGCAGCTCCGCTGTGTGTGGCTGGGAGCGGCCGAGATGCACAGCGAG 1377
 Qy 457 GlnLysLeuGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyr 476
 Db 1378 CAGATCATCGCTTCTCTCCACTACCAAGAGCGCACGGGCAATGGCAATGTGGATAC 1437
 Qy 477 GlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr 496
 Db 1438 CAGTTTGCAGTGAACCAACGACACACAGAACTACAGGTTCGGGACCTTGGAAACCAACA 1497
 Qy 497 AspTyrGluPheTyrValValAlaTyrSerGlnLysGlyAlaSerArgThrSerSerPro 516
 Db 1498 GATTATGAGTTCTACGTGTGGCTTACTCCAGCTGGGAGCGACGCGCACTCCACCCCA 1557
 Qy 517 AlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSer 536
 Db 1558 GCACCTGGTGACACACTGGATGAT-----GGTAGGGCTCTGAACCTCGCAGTGGGAGC 1611
 Qy 537 ProAsnProSerAspIleArgValAlaTrpLeuProLeuProSerLeuSerAsnGly 556
 Db 1612 -----TTGGGCTTGGACCAATGGG 1629
 Qy 557 GlnValLeuLysTyrLysLysLeuGlyTyrGlyLeuGlyLysGluAspGlnValPheSerThr 576
 Db 1630 CAGTGTGAGTACAGATAGATACGTTTGGAAAGGAGATCAGATTTTCTTACT 1689
 Qy 577 GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr 596
 Db 1690 GAGTGGAGGAAATGAGACACAGCTTATGCTGAACCTCGCTTCAGCCCAACAGGTGTAT 1749
 Qy 597 ArgValArgLysAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln 616
 Db 1750 CGAGTAGCGATTTTGGCTGTGTACAGCAGCGGCTTCCGGGGCCCCCTCCAGTGTGATGCAT 1809
 Qy 617 HisArgThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLys 636
 Db 1810 CACAGACGCCCATGTATGCAACACGAGCGATGCTCTTTCCTTTCCTGCGAGAGTTGAG 1869
 Qy 637 ValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGln 656

Db 775 ACAGAAATCCCAACCAATCAATCTTGGAGCGCTTGCATCACAATCCATGATGTC 834
Qy 287 --- AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
Db 835 TTATAACTCGGGTACTTGGAAATGGTAATCTCATGATACTGATGTCAGGCTACAACAT 894
Qy 306 SerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
Db 895 GCTGGAGTATATGTTGTCGGGCACCTACCCCTGGCACACGCACTTTACAGTTGCTATG 954
Qy 326 AlaGluLeuArgValLeuAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
Db 955 GCACTTTAACTGATTAAGTCTCTCTCAATTTGTAATGGCCAGAAATTTAAACAGG 1014
Qy 346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
Db 1015 CCTCGAGCTGGCACTGCTCGATTTGTGTGTCAGGCAAGGAATCCCTCTCCCAAGATG 1074
Qy 366 HisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGlyGly 385
Db 1075 TCATGGTTGAAATGGAAGGAAGATACATTCGAATGTTAGAAATTAATAATGTAC- 1128
Qy 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysVal 405
Db 1129 AACAGTAAATGGTAATTAACAGATTAATTCCTGAAGATGATGCTATTTATCAGTGCA 1188
Qy 406 AlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgGlu 425
Db 1189 GCTGAGNATAGCCAGATCTATTTATCTAGAGCCAGACTGACTGTAGTGTGTCAGAA 1248
Qy 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu 445
Db 1249 GACAGACCCAGTGTCTCCATTAATGTATCATGTCTGAACCATGTCAAGCTCAGCCATCT 1308
Qy 446 ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr 465
Db 1309 TTAGCTGGGAGAGGCCACTTTAATTCAGACAAAGTCAATGCCCTATTTCTGTACACTAC 1368
Qy 466 GlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAspThrThr 485
Db 1369 ATGAAAGCAAGAGTTTAATATGAAGATATCAAGTAGTCAATCGGAATATGACAACT 1428
Qy 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyr 505
Db 1429 CATATATATTATGATGACTTAGAGCTGCCAGCAATATATCTTCTACATGTAGCATAT 1488
Qy 506 SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAspVal 525
Db 1489 ATGCCAATGGAGCCAGCCAGATGTCTGACCATGTGTGACACAGAAATCTCTAGAGGATGT 1548
Qy 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
Db 1549 CCCTGAGACCTCTGAAATAGTTTGACAAAGTCGAAGTCCCAAGTCCACTGNATTTCACTCC 1608
Qy 546 TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGluTyr 565
Db 1609 TGGCTGCCAATCCAGCAATATCGCGGGGCCAAGTGTGTGTGATCGTGTCTGCTTTC 1668
Qy 566 GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
Db 1669 CGCTAGTACTGAGAAATTCACCAAGTTCTGAGCTCCCGGGGACACGCGATGATAC 1728
Qy 586 ThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGly 605
Db 1729 CTTTGGAGGCCCTGAAACCTGACAGTGTCTACTGTGTTGCGAATTAATGCTGCCACAGA 1788
Qy 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln 625
Db 1789 GTGGGCTGGAGAGTCACTAGTATGCACTTCATAGAGAGCCCAAGCT---ACAGC 1845
Qy 626 SerHisValProPheAlaProAlaGluLeuLysValArg---AlaLysMetGluSerLeu 644
Db 1845

Db 1846 GTGAAGGCCCTTAAGTCTCCA---GAGTTGCAATTTGGAGCCTCTGAACCTGTACCAACATT 1902
Qy 645 ValValSerTrpGlnProProHisProThr---GlnIleSerGlyTyrLysLeuTyr 663
Db 1903 TCTGTGAGGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTTCAGGGCTACAGCTGTAC 1962
Qy 664 TrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyArgGly 683
Db 1963 TACAAGAGAGAGGCGCAGCAGAGAAAT--- 1989
Qy 684 AspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyrGluLeu 703
Db 1990 ---GGGCCCAATTTCTTGATACCAAGCACTACTCTATATCTCTC 2031
Qy 704 ThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGlu 723
Db 2032 AGTGGCTTAGACCCCAAGAAATAATCATGTGAGACTCTCTGGCTTACAACAACATAGAC 2091
Qy 724 AspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuPro 743
Db 2092 GATGGCTATCAGGCA---GATCAGACTGTCCAGCACTCCAGGATCGGTG 2136
Qy 744 IleGlnArgGly---ProProLeuProProAlaHisValHisAlaGluSer 759
Db 2137 TCTGTTGATGATGCGATGTCCTCTCCACCAACCCACCATCTCTATGCAAGGCT 2196
Qy 760 AsnSerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIle 779
Db 2197 AACACCTCATCTTCCATCTTCTGCACTGAGGAGGCTGCTCATTCACCGCTGCACAAATC 2256
Qy 780 ValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyr 799
Db 2257 ATTAACATACACCATCGCTGTAATCTGTGTGGCTGCAGAAATGCTTCTTTGGTTCTGTAC 2316
Qy 800 TyrThrSerSerGlyGluAspIleLeuIleGlyLysLeuLysProPheThrThrTyrGlu 819
Db 2317 CTTCAACATCAGAAACTCAGATGTGTGTCAAGTCTAGAACCAACCAATATAGAA 2376
Qy 820 PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg 839
Db 2377 TTGCGCTTCGATTACATGTGGATCAGCTTCCAGTCTTCCAGCTGTGTAGTCTACCAT 2436
Qy 840 SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr 859
Db 2437 TCTACTCTTCCAGACAGCAGCAGCCACCACTGTGTGAGTAAAGTCAATTAATAGAG 2496
Qy 860 ProSerThrValArgLeuHisTyrCysProProThrGluProAsnGlyGluIleValGlu 879
Db 2497 GATGACACTGCCCTGCTGTTCTTGAAACCCCTGTGTGGCCCAAGACAGTTGTGACCCGC 2556
Qy 880 TyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuThr 899
Db 2557 TATACTATCTTATATGATCTAGGAAGGCTGGAATTCAGGAGAGTGGCAGGCTTACAC 2616
Qy 900 ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhe 919
Db 2617 CGTGAAGGGCAATAACCATGGCTTGTGTAGAAAACCTTGTAGCAGGAATGTGTACATT 2676
Qy 920 PheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
Db 2677 GTCAGATATCTGATCTCAATGAGTGGGAGGAAGGACCCCTTTTCAATTTCTGTGGAGCTG 2736
Qy 940 IleThrLeuGlnGluThrPheSerAspSer--- 949
Db 2737 GCAGTACTTCCAAAGAAACCTCTGAATCAATCAGAGGCCCAAGCGTTAGATTCTGCT 2796
Qy 950 ---LeuAspValHisAlaValThrGlyIleIle 959
Db 2797 GATGCCAAAGTTTATTTCAGGATATTACCTCGACCAAAATCAATGACTGGCATTTGCT 2856
Qy 960 ValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln 979
Db 2857 GTAGGTGTGGCATAGCCTTGATCTCATCTGTTCTCTCTCTCTATATACCGA 2916

QY	980	SerSerHisArgGluAla	-----LeuPro 987
Db	2917	AGTAAAGCCAGGAAATCATCTGCTTCCAAGCGCACAGAATGAACCTCAACAGTTACCT	
QY	988	GlyLeuSerSerSerGlyThrProGlyAsn	997
Db	2977	CGTACCAGTCGCTCTTAGCTAGTAGTGGAAAT	
RESULT 8			
ID	ABX71432		
ID	ABX71432	standard; cDNA; 3210 BP.	
XX	AC	ABX71432;	
XX	DT	06-MAR-2003 (first entry)	
XX	DE	Human cDNA encoding a novel tyrosine phosphatase, NHP1.	
XX	KW	Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;	
KW	KW	Ig super family protein; gene therapy; NHP; novel human protein;	
KW	KW	SNP; single nucleotide polymorphism.	
XX	OS	Homo sapiens.	
XX	Key	Location/Qualifiers	
FH	variation	76	
FT	FT	/*tag= a	
FT	FT	/standard_name= "Single nucleotide polymorphism"	
FT	FT	706	
FT	FT	/*tag= b	
FT	FT	/standard_name= "Single nucleotide polymorphism"	
XX	XX	US6465632-B1.	
PN	XX		
PD	XX	15-OCT-2002.	
XX	XX		
XX	XX	08-JUN-2001; 2001US-0877730.	
PF	XX		
PP	XX	09-JUN-2000; 2000US-210607P.	
PR	XX	(LEXI-) LEXICON GENETICS INC.	
PA	XX		
XX	XX	Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;	
PI	XX	Zambrowicz B, Sands AT;	
PI	XX		
XX	XX	WPI; 2003-147071/14.	
DR	XX	P-PSDB; ABUS4200.	
DR	XX		
PT	XX	Novel isolated nucleic acid which encodes a novel human protein that	
PT	XX	shares sequence similarity with animal phosphatases, that is useful for	
PT	XX	generating antibodies, and as reagents in diagnostic assays	
XX	XX		
PS	XX	Disclosure; Column 13-18; 58pp; English.	
XX	XX		
CC	XX	The invention relates to an isolated nucleic acid molecule comprising a	
CC	XX	nucleotide sequence encoding a novel human protein (NHP) that shares	
CC	XX	sequence similarity with animal phosphatases (in particular tyrosine	
CC	XX	phosphatases and is a member of the Ig superfamily). The disclosed NHP	
CC	XX	polynucleotide sequences (from genes located on either chromosome 15	
CC	XX	or chromosome 3) are useful for identifying coding sequences and in the	
CC	XX	identification of biologically relevant splice junctions. The NHP	
CC	XX	polynucleotide sequences are useful in gene therapy, for detecting	
CC	XX	mutant NHPs or inappropriately expressed NHPs, for the diagnosis of	
CC	XX	disease, for screening drugs effective in treatment of symptomatic or	
CC	XX	phenotypic manifestations of perturbing the normal function of NH.	
CC	XX	Complementary sequences of the NHP polynucleotides used in conjunction	
CC	XX	with PCR to screen libraries, isolate clones and prepare cloning and	
CC	XX	sequencing templates. Labeled NHP nucleotide probes can be used to screen	
CC	XX	a human genomic library which is helpful for identifying polymorphisms,	
CC	XX	determining the genomic structure of a given locus/allele and designing	
CC	XX	diagnostic tests. The probe sequences also have use in defining and	
CC	XX	monitoring both drug action and toxicity. Oligonucleotides	
CC	XX		

CC complementary to NHPs may encode or act as NHP antisense molecules, or
CC may be used as part of ribozyme and/or triple helix sequences.
XX The present sequence encodes an NHP.

SQ Sequence 3210 BP; 895 A; 761 C; 716 G; 836 T; 2 other;

Alignment Scores:

Pred. NO.:	5.98e-75	Length:	3210
Score:	1746.50	Matches:	397
Percent Similarity:	54.95%	Conservative:	169
Best Local Similarity:	38.54%	Mismatches:	381
Query Match:	26.45%	Indels:	83
DB:	25	Gaps:	18

US-09-754-997A-2 (1-1252) x ABX71432 (1-3210)

Qy	9	GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln	28
Db	43	GGGATGCTGCTCCGGCGCTCTGCTCTGCTGCTGCTCAGT-----CCTTGGCAGA	96
Qy	29	GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValLeuLeuGlyPro	47
Db	97	GTGTGGTGCTTTAGCGAACTGCTCTTTGTAAGAAACACACAGATGTAATGTCACAGA	156
Qy	48	GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArg	67
Db	157	AAGGACCCAGTCGTTTTAGATTTGCCAGCTCACGGAGAAGTT-----CCTATTAA	207
Qy	68	ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro	87
Db	208	GTCATGTTGAAATAATGGAGCAAAATGCTGAAATATAAATGAATGCGATCGAGTTCT	267
Qy	88	AsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAspSerAspGluGluAla	107
Db	268	AACGGCTCTTTATACATCAGT-----GAGGTGGAAGCGAGCGAGCAGCAGTCC	318
Qy	108	LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly	127
Db	319	-----GATGAAGGATTTTATCAGTGTCTGGCAATGAACAAATATGGA	360
Qy	128	ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis	147
Db	361	GCCATTCTTAGTCAAAAGCTCATCTGCTTATCACTATTATTCTGCATTTGAGTCCAG	420
Qy	148	ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly	167
Db	421	CCAATTTCCACTGAGTCCACGAAGTGGAGTTGCTCGATTTGCATGCAAGATTTCATCC	480
Qy	168	LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro	186
Db	481	CACCCCTCTGCACTATACATGGGAGTTCAATCGGACAACTCTACCTATGCTATGGAC	540
Qy	187	ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla	206
Db	541	AGGATACTGCCCTACCAACAGGAGTATTGCAATCTATGATGTACGCCAAGGATTC	600
Qy	207	GlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer	226
Db	601	GGAAATTATCGTTGATTGCTGCCACTGTAGCCCGACGCTAAAGATGATGGAGCTCG	660
Qy	227	LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIleVal	246
Db	661	CTAACTGTGATT-----CCAGTAAAGGAGTCAAAATCCTTCACACACACCAACAA	716
Qy	247	AlaAlaProGluAsnThrThrValSerGlyGlnAsnValValMetGluCysValAla	266
Db	715	GCAGGTCCACAAATACAAATCTCTTCATCAGACTGTAGTTTGGAAATGCATGGCC	774
Qy	267	SerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThr	286
Db	775	ACAGGAATCCCAACCAATCATTTCTTGGAGCGGCTTGATGACAAATCCATTGATGTC	834
Qy	287	---AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis	305

Qy 988 GlyLeuSerSerSerGlyThrProGlyValAsn 997
 Db 2977 CGTACCAGTCCCTTAGCTAGTGGAAAT 3006
 RESULT 9
 ID ABX71447 standard; cDNA; 3874 BP.
 AC ABX71447;
 XX 06-MAR-2003 (first entry)
 DT Human cDNA encoding a novel tyrosine phosphatase.
 DE Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;
 KW ig super family protein; gene therapy; NHP; novel human protein.
 XX Homo sapiens.
 OS US6465632-B1.
 PN 15-OCT-2002.
 PD 08-JUN-2001; 2001US-0877730.
 PF 09-JUN-2000; 2000US-210607P.
 PR (LEXI-) LEXICON GENETICS INC.
 PA Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;
 PI Zambrowicz B, Sands AT;
 PI WPI; 2003-147071/14.
 DR Novel isolated nucleic acid which encodes a novel human protein that
 PT shares sequence similarity with animal phosphatases, that is useful for
 PT generating antibodies, and as reagents in diagnostic assays -
 XX Disclosure; Column 107-112; 58pp; English.
 CC The invention relates to an isolated nucleic acid molecule comprising a
 CC nucleotide sequence encoding a novel human protein (NHP) that shares
 CC sequence similarity with animal phosphatases (in particular tyrosine
 CC phosphatases and is a member of the Ig superfamily). The disclosed NHP
 CC polynucleotide sequences (from genes located on either chromosome 15
 CC or chromosome 3) are useful for identifying coding sequences and in the
 CC identification of biologically relevant splice junctions. The NHP
 CC polynucleotide sequences are useful in gene therapy, for detecting
 CC mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NH.
 CC Complementary sequences of the NHP polynucleotides used in conjunction
 CC with PCR to screen libraries, isolate clones and prepare cloning and
 CC sequencing templates. Labeled NHP nucleotide probes can be used to screen
 CC a human genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests. The probe sequences also have use in defining and
 CC monitoring both drug action and toxicity. Oligonucleotides
 CC complementary to NHPs may encode or act as NHP antisense molecules, or
 CC may be used as part of ribozyme and/or triple helix sequences.
 CC The present sequence encodes an NHP.
 XX Sequence 3874 BP; 1077 A; 938 C; 897 G; 960 T; 2 other;

Alignment Scores:

Pred. No.:	7,318-75	Length:	3874
Score:	1746.50	Matches:	397
Percent Similarity:	54.95%	Conservative:	169
Best Local Similarity:	38.54%	Mismatches:	381
Query Match:	26.45%	Indels:	83
DB:	25	Gaps:	18

US-09-754-997A-2 (1-1252) x ABX71447 (1-3874)
 Qy 9 GlyLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln 28
 Db 233 GGGAGTCTGCTCCGGCGCTCTGCTCTGCTGCTGCTCAGT-----CCTTGGCAGA 286
 Qy 29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValLeuLeuGlyPro 47
 Db 287 GTGTGTGTCTTTAGCGAAGCTGCTTTTGTAAAGAACACACAGAGTGTAACTGTCAACA 346
 Qy 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProThrArg 67
 Db 347 AAGGACCCAGTGTGTTTAGATTGCCAGGCTCAGGAGAAAGTT-----CCTATTAA 397
 Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuLeuPro 87
 Db 398 GTCACATGTTGAAAATGAGCAAAAATGCTGAAAATAAACAGATCAGGTCTCTTCT 457
 Qy 88 AsnGlySerLeuTrpLeuSerProLeuGluGlnGluAspSerAspAspGluAla 107
 Db 458 AACGGCTCTTTATACATCAGT-----GAGGTGGAAGCGAGCGAGAGCAGTCC 508
 Qy 108 LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly 127
 Db 509 -----GATGAAGGATTTTATCAGTGTCTGCAATGAAACAATATGGA 550
 Qy 128 ValValAlaSerGlnValAlaValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
 Db 551 GCCATTCTTAGTCAAAAGCTCATCTTGCTTATCAACTATTTCTGCATTTGAGTCCAG 610
 Qy 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
 Db 611 CCAATTTCCACTGAGGTCCACGAAGGTGGAGTTGCTCGATTTGTCATGCAAGATTTTCATCC 670
 Qy 168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
 Db 671 CACCCTCTCGAGTCATAACATGGGAGTTCAATCGGACAACTCTACCTATGATATGAC 730
 Qy 187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
 Db 731 AGATTAACCTGCCCTACCAACAGGAGTATTCAGATCTATGATGTGACCCAAAGGATCT 790
 Qy 207 GlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
 Db 791 GGAATATCTGTTGTTATGCTGCCACTGTAGCCACCGACCGATAAAGATATGAGGCGCTCG 850
 Qy 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValIleVal 246
 Db 851 CTAACCTGTGATT-----CCAGCTAAGAGTCAAAATCTCTCCACACACACCAATATA 904
 Qy 247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
 Db 905 GCAGGTCCACAGAACATAACACATCTCTTCATCAGACTGATGTTTGGATATGATGCC 964
 Qy 267 SerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThr 286
 Db 965 ACAGGAATCCCAACCAATCAATTTCTTGGAGCGCTTGCATCACAATCCATTGATGTC 1024
 Qy 287 ---AspValIleValLeuGlyArgThrAsnLeuLeuAlaSerAlaGlnProArgHis 305
 Db 1025 TTTAATACTCGGGTACTTGGAAATGGTAATCTCATGATATCTGATGTCAGCGTACAACAT 1084
 Qy 306 SerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
 Db 1085 GCTGAGTATATGTTTGTGCGGCACCTACCCCTGGCACAGCAACTTACAGTTGCTATG 1144
 Qy 326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
 Db 1145 GCAACTTTAAGTATGATGCTCTCTCTTCAATTTGATGGCCAGAAAGTTTAAACAGG 1204
 Qy 346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
 Db 1205 CCTCGAGCTGGCACTGCTCGATTGTTGTGTGTCAGGACGAAGAAATCCCTCTCTCCCAAGATG 1264

QY 366 HisTrpLeuHisaspGlyIleProLeuArgProAsnGlyArgValIysValGlnGlyGly 385
DB 1265 TCATGGTTGAAATGGAAGCAATACATTCGAATCGTAGAATTTAAATGTATC----- 1318
QY 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrGlnCysVal 405
DB 1319 AACAGTAAATGGTAATTAACAGATTAATTCCTGAAGATGATGCTATTATATCAGTGCATG 1378
QY 406 AlaGlnAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgGlu 425
DB 1379 GCTGAGAATAGCAAGGATCTATTATCTAGAGCCAGACTGACTGATGATGATGCAGAA 1438
QY 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu 445
DB 1439 GACAGACCCAGTGTCTCCCTAATATGATCATCTGCTGAACCACTGTCAAGCTCAGCCATCTT 1498
QY 446 ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr 465
DB 1499 TTAGCCTGGGAGAGGCCACTTTATATTCAGACAAGTCAATGCTATTCTGTACACTAC 1558
QY 466 GlnIysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThr 485
DB 1559 ATGAAGCAGAGAGGTTAATATGAAGATGATCACTAGTAGTCATCGAATATGACACACT 1618
QY 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyr 505
DB 1619 CATTTATATTATGATGACTTAGAGCCTGCCAGCAATATATCTTCTACATGTAGCATAT 1678
QY 506 SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAspVal 525
DB 1679 ATGCCAATGGAGCCAGCCAGATGTCTGACCATGTGTGACACAGAAATCTCTAGAGGATGT 1738
QY 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
DB 1739 CCCTGAGACTCTGAAATAGTTTGACAGTCCAGTCCAGTCCACTGATATCTCATCTCC 1798
QY 546 TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuIysTyrIleGluTyr 565
DB 1799 TGGCTGCCAATCCAGCCAAATATCGCGGGGCCAAGTGTGTGTATCGTCTGTCTTCTTTC 1858
QY 566 GlyLeuGlyIysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
DB 1859 CGCTAATGATCTAGAAATCAATCCAAAGTTCTGAGAGCTCCCGGGACACGCAATGATAC 1918
QY 586 ThrLeuAsnSerLeuGlnProAsnValValTyrArgValArgIleSerAlaGlyThrGly 605
DB 1919 CTTTGGAGCCCTGAAGTCTGACAGTGTCTACTGTTCCGATTAATCTGCTGCCACAGA 1978
QY 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln 625
DB 1979 GTGGGCTGGAGAGTCAATCAGTATGCACTTCACATAGAGCCCAAGCT---ACAAGC 2035
QY 626 SerHisValProPheAlaProAlaGluLeuIysValArg---AlaIysMetGluSerLeu 644
DB 2036 GTCAAGCCCTAAGTCTCA---GAGTTGCAATTTGGAGCCTCTGAACTGTACCAACCAT 2092
QY 645 ValValSerTrpGlnProProHisProThr---GlnIleSerGlyTyrIysLeuTyr 663
DB 2093 TCTGTAGGTGGCAGCAAGATGATAGAGGACACAGCTCTATCAGGCTTCAAGCTGTATC 2152
QY 664 TrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyArgGly 683
DB 2153 TACAAGGAGAGGGCAGGAGAAAT----- 2179
QY 684 AspGlnAlaTrpAspValGlyProValArgLeuIysLysValIysGlnTyrGluLeu 703
DB 2180 -----GGGCCCAATTTCTTGGATACCAAGGACCTACTCTATACTCTC 2221
QY 704 ThrGlnLeuValProGlyArgProTyrGluValIysLeuValAlaPheAsnLysHisGlu 723
DB 2222 AGTGGCTTAGACCCCAAGAGAAATATCATGTGAGACTCTGCGCTTCAACAACATAGAC 2281

QY 724 AspGlyTyrAlaAlaValTrpIysGlyIysThrGluIysAlaProThrProAspLeuPro 743
DB 2282 GATGGCTATCAGGCA-----GATCAGACTGTCCAGCACTCCAGGATCGGTG 2326
QY 744 IleGlnArgGly-----ProLeuProProAlaHisValHisAlaGluSer 759
DB 2327 TCTGTTGATGCGATGCTCTCTCTCCACACACCCACCATCTCTATGCGAAGCT 2386
QY 760 AsnSerSerThrSerIleTrpLeuArgTrpIysLysProAspPheThrThrValIysIle 779
DB 2387 AACACCTCATCTTCCATCTTCTGCACTGGAGAGGCTGCATTCAACCGTCGCAAAATC 2446
QY 780 ValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyr 799
DB 2447 ATTAACATACCATCGCTGTAATCTGTGGCTGCAGATGCTTCTTGGTCTGTATC 2506
QY 800 TyrThrSerSerGlyGluAspIleLeuIleGlyGlyLeuIysProPheThrIysTyrGlu 819
DB 2507 CTTCAACATCAGAACTCAGATGTGGTTCAGGTCTAGAACCAACCAACCAATACGAA 2566
QY 820 PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg 839
DB 2567 TTGGCGTTCCGATTACATGTCATCAGCTTTCAGTCTTGGAGCCCTGTAGTCTACCAT 2626
QY 840 SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr 859
DB 2627 TCTACTCTTCCAGAACACACAGCCAGCCAGCTTGGAGTAAAGTGACATTAATAGAG 2686
QY 860 ProSerThrValArgLeuHisTrpCysProThrGluProAsnGlyGluIleValGlu 879
DB 2687 GATGACACTGCCCTGTTCTTGGAAACCCCTGTATGGCCCAAGAACAGTTGGACCCG 2746
QY 880 TyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThr 899
DB 2747 TATACTATCTATATGCACTAGGAAGGCTGATTCAGAGAGAGTGGCAGGCTTTACAC 2806
QY 900 ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhe 919
DB 2807 CGTGAAGGGCAATACCATGCTTGGCTAGAAAACCTTGTGTAGCAGGAAATGTGTACAT 2866
QY 920 PheIysMetGlyValArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
DB 2867 GTCAAGATATCTGCATCCCAATGAGGTGGGAGGAGGACCTTTTCAAATTTCTGTGAGCTG 2926
QY 940 IleThrLeuGlnGluThrPheSerAspSer----- 949
DB 2927 GCAGTACTTCCAAAGGAACCTCTGAATCAAATCAGAGCCCAAGCGTTTAGATTCTGCT 2986
QY 950 -----LeuAspValHisAlaValThrGlyIleIle 959
DB 2987 GATGCCAAAGTTTATTTCAGATATTACCATCTGCACCAAAAATCAATGACTGCAATGCT 3046
QY 960 ValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln 979
DB 3047 GTAGGTGTGGCATACGCTTGACCTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3106
QY 980 SerSerHisArgGluAla-----LeuPro 987
DB 3107 AGTAAGCCAGGAATATCTGCTTCCAGAGCCGACAGATGGAATCAACTCAAGTTTACCT 3166
QY 988 GlyLeuSerSerSerGlyThrProGlyAsn 997
DB 3167 CGTACCAGTGCCTCTCTTAGCTAGTGGAAAT 3196

RESULT 10

ABX71437

ID ABX71437 standard; cDNA; 2976 BP.

XX AC

XX ABX71437;

XX 06-MAR-2003 (first entry)

XX DT

XX Human cDNA encoding a novel tyrosine phosphatase, NHP6.

CC nucleotide sequence encoding a novel human protein (NHP) that shares
 CC sequence similarity with animal phosphatases (in particular tyrosine
 CC phosphatases and is a member of the Ig superfamily). The disclosed NHP
 CC polynucleotide sequences (from genes located on either chromosome 15
 CC or chromosome 3) are useful for identifying coding sequences and in the
 CC identification of biologically relevant splice junctions. The NHP
 CC polynucleotide sequences are useful in gene therapy, for detecting
 CC mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NHP.
 CC Complementary sequences of the NHP polynucleotides used in conjunction
 CC with PCR to screen libraries, isolate clones and prepare cloning and
 CC sequencing templates. Labeled NHP nucleotide probes can be used to screen
 CC a human genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests. The probe sequences also have use in defining and
 CC monitoring both drug action and toxicity. Oligonucleotides
 CC complementary to NHPs may encode or act as NHP antisense molecules, or
 CC may be used as part of ribozyme and/or triple helix sequences.
 CC The present sequence encodes an NHP.

XX SQ Sequence 3219 BP; 927 A; 757 C; 706 G; 828 T; 1 other;

Alignment Scores:

Pred. No.: 2,61e-72 Length: 3219
 Score: 1691.50 Matches: 375
 Percent Similarity: 55.79% Conservative: 160
 Best Local Similarity: 39.10% Mismatches: 347
 Query Match: 25.61% Indels: 77
 DB: 25 Gaps: 15

US-09-754-997A-2 (1-1252) x ABX71440 (1-3219)

QY 79 GluHisGluAsnLeuHisLeuProAnGlySerLeuTrpLeuSerSerProLeuGlu 98
 Db 7 GAAATTAACGATCGAGGTTCTTTCTAAAGGCTCTTTATACATCAGT-----GAG 57
 QY 99 GlnGluAspSerAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTyr 118
 Db 58 GTGGAGGCGCGAGGAGGAGTCC-----GATGAGGATTTTAT 99
 QY 119 SerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValLysLeu 138
 Db 100 CAGTCTGGCAATGAACAATATGAGGCAATCTTAGTCAAAAGCTCATCTCCCTTA 159
 QY 139 AlaThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThr 158
 Db 160 TCAACTATTCTGCAATTTGAAGTCCAGGCCAATTTCCACTGAGGTCCACGAAAGGTGAGTT 219
 QY 159 AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAsp 178
 Db 220 GCTCGATTGCGATGCAAGATTTCATCCACCCTCTCGAGTCATTAACATGGAGTTCAAT 279
 QY 179 GlnValThrValPro---GluGluProArgLeuIleThrLeuProLysTrpLeuGluGln 197
 Db 280 CGGACAACTCTACCTATGACTATGACAGGATAACTGCCCTACCAACAGGAGTATTCGAG 339
 QY 198 IleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAla 217
 Db 340 ATCTATGATGTGAGCCAAAGGAGTTCTGGAAATTTATCGTTGTTGCTGCCACTGTAGCC 399
 QY 218 ArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAla 237
 Db 400 CACCGACGTAAAGTATGGAGCGCTCGCTAACTGTGATT-----CCAGCTAAGGAGTCA 453
 QY 238 ThrArgGlyGlnAspValIleValAlaAlaProGluAsnThrThrValValSerGly 257
 Db 454 AAATCTCTCCACACACCAACCAATATGACAGTCCACAGAACATAACACATCTCTTCAT 513
 QY 258 GlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpVal 277
 Db 514 CAGACTGTAGTTTGGAAATGATGGCCACAGGAAATCCCAACCAATCATTTCTTTGGAGC 573

QY 278 ArgGlnAspGlyLysProLysSerThr---AspValIleValLeuGlyArgThrAsnLeu 296
 Db 574 CGCCTTGATCACAATCCATCTGATGCTTTAAATACCTCGGGTACTTGGAAATGTAATCTC 633
 QY 297 LeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysPro 316
 Db 634 ATGATATCTCATGTCAGGCTACCAACATGCTGGAGTATATGTTGTGGGCGCACTACCCCT 693
 QY 317 LeuThrArgAspPheAlaThrAlaAlaLeuArgValLeuAlaAlaProAlaIle 336
 Db 694 GGCACACGCACTTACAGTTGCTATGGCACTTTAACTGCTATTAGCTCCTCTTCATTT 753
 QY 337 SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg 356
 Db 754 GTTGAATGGCCAGAAAGTTTAAACAGGCTCGAGCTGGCACTGCTGCAATTTGTGTGTCAG 813
 QY 357 AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyLysLeuProLeuArgPro 376
 Db 814 GCGAAGGAATCCCTCTCCCAAGATGTCATGTTGAAAATGGAAGGAAGATCATTCG 873
 QY 377 AsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeu 396
 Db 874 AATGTTAGATTAATAATGTAC-----AACAGTAAATTTGTTAATTAACCAATATTATCTCT 927
 QY 397 GlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAla 416
 Db 928 GAAGATGATGCTATTTATCAGTGCATGCTGAGATAGCAAGGATCTATTATTATCTAGA 987
 QY 417 AlaProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThrAla 436
 Db 988 GCCAGACTGACTGTAGTGTGTGAGACAGACAGACCCAGTCTCCCTATAATATGAAGATAT 1047
 QY 437 ThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlu 456
 Db 1048 GAAACCATGTCAGCTCAGCCATCTTTTAGCTGGAGAGGCCACTTTTATAATTCAGAC 1107
 QY 457 GlnIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyr 476
 Db 1108 AAGTCATTGCTATTCTGTACACTACATGAAGACAGAGGTTTAAATAATGAAGAGTAT 1167
 QY 477 GlnPheAlaValAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr 496
 Db 1168 CAAGTAGTCATCGAAATGACACAACTCATTTATATTATGATGACTTAGAGCTGCGCAGC 1227
 QY 497 AspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerPro 516
 Db 1228 AATTATCTTCTACATTTAGCATATATGCCAATGGAGGCCAGCCAGATCTCTGACCAT 1287
 QY 517 AlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSer 536
 Db 1288 GTGACACAGAATACTCTAGAGGATGTTCCCTCGAGACTCTCTGAAATTAGTTTGACAAGT 1347
 QY 537 ProAsnProSerAspIleArgValAlaTrpLeuProLeuProLeuProSerSerLeuSerAsnGly 556
 Db 1348 CGAATGCCACTGATATTTCTCATCTCTGCTGCTGCAATCCAGCCAAATATCGCGGGGCG 1407
 QY 557 GlnValLeuLysTyrLysIleGluTyrGlyLysGluAspGlnValPheSerThr 576
 Db 1408 CAAGTGGTGTGATTCGCTGCTCTTCGCTTAGTACTGAGAAATTCATCAAGTTCTG 1467
 QY 577 GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr 596
 Db 1468 GAGCTCCCGGGGACCCAGCATGAGTACCTTTTGGAGGCCCTGAAACCTGACAGTGTCTAC 1527
 QY 597 ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln 616
 Db 1528 CTGGTTTCGATTTACTGCTGCCACAGTGGCGCTGGAGAGTCAATCAGTAGTGGACTTCA 1587
 QY 617 HisArgThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLys 636
 Db 1588 CATAGGACGCCCAAGCT---ACAAGCGTGAAGGCCCTAAGTCTCCA---GAGTTGAT 1641
 QY 637 ValArg---AlaLysMetGluSerLeuValValSerTrpGlnProProProHisProThr 655

Db	1642	TTGGAGCCTCTGAAC	TGTA	CAACCATTTCTGTGAGG	TGGCAGCAAGATGTAGAGG	CAACA	1701
Qy	656	---Gln	IleSerGly	TyrLys	LeuTyrTrp	GlyGluValGlyThrGluGluGluAla	Asp 674
Db	1702	GCTGCTATT	CAGGCTCA	AGCTGTACTACA	GGAAGGAGG	CAGCAGAGAGAT	-----1755
Qy	675	GlyAsp	ArgPro	ProGlyGly	ArgGlyAsp	GlnAlaTrpAspValGlyProVal	ArgLeu 694
Db	1756	-----	-----	-----	-----	GGGCCCAATTTCTTG	1770
Qy	695	LysLys	LysValLys	GlnTyrGlu	LeuThrGln	LeuValProGlyArgProTyrGlu	Val 714
Db	1771	GATACCA	AGGACCTACT	CTATAC	TCTCAGTGGCTTAG	ACCCAGAGAAATATCAT	GTG 1830
Qy	715	LysLeu	ValAla	LeuPhe	AsnLys	HisGluAspGlyTyrAla	ValTrpLysGlyLysThr 734
Db	1831	AGACT	CTCTGGCTT	ACAACA	CATAGACG	TATCAGCA	-----GAT 1875
Qy	735	GluLys	AlaPro	ThrPro	AspLeu	ProIleGlnArgGly	-----ProProLeu 750
Db	1876	CAGACTGT	CAGCACTCC	GAGATGGT	GTCTGTCG	TATGCATGGTCCCTCTCC	CAACA 1935
Qy	751	ProPro	AlaHis	ValHis	AlaGlu	SerAsnSerThrSer	IleTrpLeuArgTrpLys 770
Db	1936	CCACCC	CACCATCTCT	ATGCAAGGCTTA	ACACACCTCATCTT	CCATCTTCTCCGCACTGGAGG	1995
Qy	771	LysPro	AspPhe	ThrThr	ValLys	IleValAsnTyrThrVal	ArgPheGlyProTyrGly 790
Db	1996	AGGCTG	CATTCACCG	TGTCACA	ATCATTA	CTACACCATCCGCTGTATCTCTGT	TGGC 2055
Qy	791	LeuArg	AsnAla	SerLeu	ValThrTyrTyrThr	SerSerGlyGluAsp	IleLeuIleGly 810
Db	2056	CTCGA	ATGCTCTCTT	GGTTCTGTAC	CTTCAAA	CATCAGAACTCATGTTGGT	TCAA 2115
Qy	811	GlyLeu	LysPro	PheThrLys	TyrGlu	PheAlaValGlnSerHis	GlyValAspMetAsp 830
Db	2116	GGCTAG	AAACCAAC	ACCAATACGA	ATTTGCCG	TTCGATTACATGTG	GCATCAGCTTTCC 2175
Qy	831	GlyPro	PheGlySer	ValVal	GluArgSerThr	LeuProAspArgPro	SerThrProPro 850
Db	2176	AGTCCT	TGGAGCCCTGT	AGTCTAC	CTTCTCTTCC	AGAGCACGACGAGCC	CCACCA 2235
Qy	851	SerAsp	LeuArg	LeuSerPro	LeuThrPro	SerThrValArgLeuHis	TyrCysProPro 870
Db	2236	GTTGG	AGTAAAAGTG	ACATTA	TAGAGGATG	ACACTGCCCTGGT	TTCTTGGAAACCCCT 2295
Qy	871	ThrGlu	ProAsnGly	GluIleVal	GluTyrLeu	IleLeuTyrSerAsn	AsnHisThrGln 890
Db	2296	GATGG	CCAGAAAC	AGTTGTG	ACCCGCTAT	ACTATCTATATG	CATCTAGGAAGCCCTGG 2355
Qy	891	ProGlu	HisGln	TrpThrLeu	LeuThrThr	GluGlyAsnIlePhe	SerAlaGluValHis 910
Db	2356	ATTC	CAGAGAGTGG	CAGGCTCTT	ACACCGTGA	GGGGCAATAACCATGG	CTTTGCTAGAA 2415
Qy	911	GlyLeu	GluLys	AspThrArg	TyrPhe	PheLysMetGlyAla	ArgThrGluValGlyPro 930
Db	2416	AACTT	GGTAGCAG	GAATGTCT	CAATGTCA	GATATCTGCATC	CAATGAGTGGGAGAA 2475
Qy	931	GlyPro	PheSerArg	LeuGlnAsp	ValIleThrLeu	GlnGlnThrPhe	SerAspSer---949
Db	2476	GGAC	CCCTTTTCA	ATCTCTGT	GGAGCTGG	CAGTACTTCCAA	AGGAAACCTCTGAATCAAAT 2535
Qy	950	-----	-----	-----	-----	-----	Leu 950
Db	2536	CAGAG	CCCCAAG	CGTTAG	ATCTCTGT	GATGCCAAAGTTTAT	TATCAGGATATTACCATCTG 2595
Qy	951	AspVal	HisAla	ValThrGly	IleIleVal	GlyValCysLeu	GlyLeuLeuCysLeuLeu 970
Db	2596	GACCA	AAAAATCA	TGACTG	GCATTTG	TGTAGGTGG	CATAGCCTTGACCTGCATCCTC 2655
Qy	971	AlaCys	MetCys	AlaGlyLeu	ArgGlnSer	SerHisArg	GluAla-----985

[illegible]

SQ Sequence 825 BP; 174 A; 271 C; 236 G; 144 T; 0 other;

Alignment Scores:
Pred. No.: 2,06e-62 Length: 825
Score: 1472.00 Matches: 275
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.29% Indels: 0
DB: 22 Gaps: 0

US-09-754-997A-2 (1-1252) x AADI0023 (1-825)

```
QY 978 ArgGlnSerSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAen 997
Db 1 CGCAAAAGCTCCACAGGAAGCCCTTCCCGATTGTCTCTCAGGCACCCAGGAAC 60

QY 998 ProAlaLeuTyThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeu 1017
Db 61 CCAGCGCTCTACACAAAGAGCTCGGCTTGGGCTCTCCAGTGCTCCCTGCTGCCCATGAGTTG 120

QY 1018 GluSerLeuValHisProArgProGlnAspTpsrProProSerAspValGluAsp 1037
Db 121 GAGTCCCTCGTGATCTCTGTCCTCCAGGATGGTCCCAACCCCTCAGATGTGGAAGAC 180

QY 1038 LysAlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerIys 1057
Db 181 AAGGCTGAAGTACACAGCCCTTATGGTGGCAGTGTTCAGATTGCCGGGCGCACTCCCAAG 240

QY 1058 ArgIysIleSerTrpAlaGlnAlaGlyProAsnTrpAlaGlySerTrpAlaGlyCys 1077
Db 241 AGAAGATCTCTGGGCTCAGCGAGGGGAGCAAACTGGGCGAGGCTCTGGGCGAGGCTGT 300

QY 1078 GluLeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProAla 1097
Db 301 GAGTGGCCCAAGGTAGTGTGTCACAGCGCGCTCTGACCCCTGTCTGCTGCTCCAGCG 360

QY 1098 GlyThrGlyGlnThrLeuLeuLeuGlnAlaLeuValTyArgGlyIleLysSerAenGly 1117
Db 361 GGAACCGGGCAGACACTCTGCTGCAAGCCCTGTGTATGACGGCATAAAGAGCAACGGG 420

QY 1118 ArgIysIleProSerProAlaCysArgAenGlnValGluAlaGluValIleValHisSer 1137
Db 421 AGAAGAGAGCGTCTCCCAAGCTGCGAGGAATCAGGTGGAAGCTGAGGTCAATGTCCATCC 480

QY 1138 AspPheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGlu 1157
Db 481 GACTCGGTGATCCAAAGAGATGCTCTGACCTCCACCTCCAGACCTGGAGCCAGAGGAA 540

QY 1158 ProLeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAla 1177
Db 541 CCACTGACTGCAGAGACTCTGCTTCCACGCTGCGAGCTGTGGATCTGTCTCAAGGAGCA 600

QY 1178 AspTrpLeuGlyArgGluLeuGlyClyCysGlnProThrThrSerGlyProGluArgLeu 1197
Db 601 GACTGGCTGGGCGAGGAGCTGGAGGGTGGCAACCAACACAGTCAGTGGGCGCAGAGGGCTC 660

QY 1198 ThrCysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThr 1217
Db 661 ACCTGCTTGCACAGAGCAGCAGTGCCTCTGCTCTGCTCTCAGACCTCCACCCAGCACT 720

QY 1218 AlaIleGluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrVal 1237
Db 721 GCTATAGAGGAGGCCCTGGGAAAGCTGCCAGGCCCAAGCCCTGTGTCTCTTAACATGC 780

QY 1238 SerProSerLeuProArgAlaProValSerSerAlaGlnValPro 1252
Db 781 AGCCCAAGCCTTCCAGGGGCCCTGTCTCTCTGTCTGCTCAGGTCCCC 825
```

RESULT 13

ID AAS76658

XX AAS76658 standard; cDNA; 3330 BP.

AC AAS76658;

```
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #12462.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; es.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSBQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG12471.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX Claim 1; SEQ ID No 12462; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 3330 BP; 971 A; 775 C; 746 G; 838 T; 0 other;

Alignment Scores:
Pred. No.: 4,05e-61 Length: 3330
Score: 1458.50 Matches: 341
Percent Similarity: 50.58% Conservatives: 139
Best Local Similarity: 35.93% Mismatches: 324
Query Match: 22.09% Indels: 145
DB: 23 Gaps: 16

US-09-754-997A-2 (1-1252) x AAS76658 (1-3330)
QY 131 SerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHisProGluSer 150
Db 226 AGTAGCTCAATGTTAGTGAACACCACTTCTATTCTGCAATTCAGTCCAGCCATTTCC 285
QY 151 GlnIleValGluGluAenGlyThrAlaArgPheGluCysHisThrIysGlyLeuProAla 170
```


QY 187 ArgLeuIleThrLeuProIleThrLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
| | | | |
Db 541 AGATAACTGCCCTCAACAGGAGTATTCAGATCTATGATGTCAGCCAAAGGGATTCT 600
| | | | |
QY 207 GlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
| | | | |
Db 601 GGAATATTCGTGTATTCCTGCACTGTAGCCACCGAGCTAAAGTATGGAGGCTCG 660
| | | | |
QY 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIleVal 246
| | | | |
Db 661 CTAACGTGTATT-----CCAGCTAAGGAGTCAAAATCCTCCACACACCAACCAATTATA 714
| | | | |
QY 247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
| | | | |
Db 715 GCAGGTCCACAGAACATAACAACATCTCTTCATCAGACTGTAGTTTTGGAAATGCATGGCC 774
| | | | |
QY 267 SerAlaAspProThrProPheValSerTyrValArgGlnAspGlyIlePheProIleSerThr 286
| | | | |
Db 775 ACAGGAATCCCAACCAATCATTTCTTGAGCGCGCTTGATCACAATCCATTGATGTC 834
| | | | |
QY 287 --- AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
| | | | |
Db 835 TTTAATACTCGGTACTTGGAAATGTAATCTCATGTATCTGATGTGAGGTACAACT 894
| | | | |
QY 306 SerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
| | | | |
Db 895 GCTGGAGTATATGTTGTGGGCCACTACCCCTGGCACAGCACTTTACAGTTGCTATG 954
| | | | |
QY 326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
| | | | |
Db 955 GCACCTTTAATCTATGATCTCTCTCTTCAATTTGTAATGGCCAGAAAGTTTAAACAGG 1014
| | | | |
QY 346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlnProArgProAlaLeu 365
| | | | |
Db 1015 CCTCGAGCTGGCACTGCTCGATTTGTGTGTCAGGAGAGAAATCCCTCTCCCAAGATG 1074
| | | | |
QY 366 HisTyrLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGlyGly 385
| | | | |
Db 1075 TCATGTTGAAAATCGAAGAGATACATTCGAATGGTAGAATTAATATGTAC----- 1128
| | | | |
QY 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysVal 405
| | | | |
Db 1129 AACAGTAATGTTGTAATTAACAGATTTATCTCGAAGATGATGCTATTTATCAGTGCATG 1188
| | | | |
QY 406 AlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgGlu 425
| | | | |
Db 1189 GCTGAGATAGCCAGAGATCTATTTATCTAGAGCCAGCTGACTGTAGTATGTCAGAA 1248
| | | | |
QY 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu 445
| | | | |
Db 1249 GACAGACCCAGTCTCCCTATAATGTATCATGTCTGAAACCATGTCAAGCTCAGCCATTCTT 1308
| | | | |
QY 446 ValAlaTyrGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr 465
| | | | |
Db 1309 TTAGCCTGGGAGAGGCCACTTTATAATTTCAGAAAGTCATTTGCTGTACACTAC 1368
| | | | |
QY 466 GlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThr 485
| | | | |
Db 1369 ATGAAAGCAGAGAGTTTAAATAATGAAGAGATCAAGTAGTCTACGGAATGACACACT 1428
| | | | |
QY 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyr 505
| | | | |
Db 1429 CATTATATTATGATGACTTAGAGCTGCCAGCAATATATCTTTCTACATTGTAGATAT 1488
| | | | |
QY 506 SerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHisThrLeuAspAspVal 525
| | | | |
Db 1489 ATGCCAATGGAGCCAGCAGATGCTGACCATGTGACACAGAAATCTCTAGAGGATGAC 1548
| | | | |
QY 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
| | | | |
Db 1549 CCAGA----- 1554

QY 546 TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrIleGluTyr 565
| | | | |
Db 1554 ----- 1554
| | | | |
QY 566 GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
| | | | |
Db 1554 ----- 1554
| | | | |
QY 586 ThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGly 605
| | | | |
Db 1554 ----- 1554
| | | | |
QY 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln 625
| | | | |
Db 1554 ----- 1554
| | | | |
QY 626 SerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuVal 645
| | | | |
Db 1554 ----- 1554
| | | | |
QY 646 ValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrpGly 665
| | | | |
Db 1554 ----- 1554
| | | | |
QY 666 GluValGlyThrGluGluAlaAspGlyAspArgProProGlyArgGlyAspGln 685
| | | | |
Db 1554 ----- 1554
| | | | |
QY 686 AlaTrpAspValGlyProValArgLeuLysLysValLysValGlnTyrGluLeuThrGln 705
| | | | |
Db 1554 ----- 1554
| | | | |
QY 706 LeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGluAspGly 725
| | | | |
Db 1555 -----AGAAATATCATGTGAGATCTCTGGCTTACACACATACACATGAC 1602
| | | | |
QY 726 TyrAlaAlaValTrpLysGlyThrGluLysAlaProThrProAspLeuProIleGln 745
| | | | |
Db 1603 TATCAGGCA-----GATCAGACTGTGAGCACTCCAGGATGCGTGTCTGT 1647
| | | | |
QY 746 ArgGly-----ProLeuProProAlaHisValHisAlaGluSerAsnSer 761
| | | | |
Db 1648 CGTGATCGCATGTCCTCTCCACACACCCCACTCTCTATGCGAAGCTAACCA 1707
| | | | |
QY 762 SerThrSerIleTrpLeuArgTyrLysLysProAspPheThrThrValLysIleValAsn 781
| | | | |
Db 1708 TCATCTTCATCTCTCTGACCTGGAGAGGCTGCAATTCACCGCTGCACAAATCATTAAC 1767
| | | | |
QY 782 TyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThrTyrTyrThr 801
| | | | |
Db 1768 TACACCATCGCTGTAATCTGTCCTGCGCTGCAGATGCTCTCTTGTCTGTACCTCAA 1827
| | | | |
QY 802 SerSerGlyGluAspIleLeuIleGlyGlyLeuLysProPheThrLysTyrGluPheAla 821
| | | | |
Db 1828 ACATCAGAAACTCATCATGTGGTCTCAAGGCTTAGAACCAACACCAAAATACGAAATTCGC 1887
| | | | |
QY 822 ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr 841
| | | | |
Db 1888 GTTCATTATCATGTGATCAGCTTCCAGTCTCTGAGCCCTGTAGTCTTACCATCTACT 1947
| | | | |
QY 842 LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer 861
| | | | |
Db 1948 CTTCCAGAACACACAGCCGCCACCATGTTGGAGTAAAGTCATTAATAGAGGATGAC 2007
| | | | |
QY 862 ThrValArgLeuHisTyrCysProProThrGluProAsnGlyGluIleValGluTyrLeu 881
| | | | |
Db 2008 ACTGCCCTGGTCTCTTGGAAACCCCTGATGCCCCAGAAACAGATGTGACCGCTATACT 2067
| | | | |
QY 882 IleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeuThrThrGlu 901
| | | | |
Db 2068 ATCTTATATGATCTAGGAGGCCCTGGATTCGAGGAGAGTGGCAGGTCTTACACCGTGAA 2127
| | | | |
QY 902 GlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLys 921
| | | | |

Db 2128 GGGCAATACCATGCTTGTAGTAACTTGTAGCAGGAATGTGTACATGTCAAG 2187
 Qy 922 MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
 Db 2188 ATATCTGCATCCAAATGAGTGGGAGAGGACCCCTTTTCAAAATCTGTGGAGCTGGCAGTA 2247
 Qy 942 LeuGlnGluThrPheSerAspSer----- 949
 Db 2248 CTTCAAGGAAGAACCTTGAAATCAATCAGAGGCCCAAGCGTTTAGATTTCTGTGTATGCC 2307
 Qy 950 -----LeuAspValHisAlaValThrGlyIleIleValGly 961
 Db 2308 AAAGTTTATTACCATTTACCATCTGGACCAAAATCAATGACTGGCATTGTCTGTAGT 2367
 Qy 962 ValCysSerGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSerSer 981
 Db 2368 GTTGGCATAGCTTTGACCTGCATCTCATCTGTGTCTTCATCTTTGATATACCGAAGTAAA 2427
 Qy 982 HisArgGluAla-----LeuProGlyLeu 989
 Db 2428 CCGAGGAATCATCTGCTTCCAGACGCCACAGATGGAATCAACAGTTACCTCGTACC 2487
 Qy 990 SerSerSerGlyThrProGlyAsn 997
 Db 2488 AGTGCTCTTGTAGCTAGTGGAAAT 2511
 RESULT 15
 ID ABX71436 standard; cDNA; 2958 BP.
 AC ABX71436;
 DT 06-MAR-2003 (first entry)
 DE Human cDNA encoding a novel tyrosine phosphatase, NHP5.
 KW Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;
 KW Ig super family protein; gene therapy; NHP; novel human protein.
 OS Homo sapiens.
 PN US6465632-B1.
 PD 15-OCT-2002.
 PF 08-JUN-2001; 2001US-0877730.
 PR 09-JUN-2000; 2000US-210607P.
 PA (LEXI-) LEXICON GENETICS INC.
 PI Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;
 PI Zambrowicz B, Sands AT;
 DR WPI; 2003-147071/14.
 DR P-PSDB; ABUS4204.
 PT Novel isolated nucleic acid which encodes a novel human protein that
 PT shares sequence similarity with animal phosphatases, that is useful for
 PT generating antibodies, and as reagents in diagnostic assays -
 PS Disclosure; Column 41-44; 58pp; English.
 CC The invention relates to an isolated nucleic acid molecule comprising a
 CC nucleotide sequence encoding a novel human protein (NHP) that shares
 CC sequence similarity with animal phosphatases (in particular tyrosine
 CC phosphatases and is a member of the Ig superfamily). The disclosed NHP
 CC polynucleotide sequences (from genes located on either chromosome 15
 CC or chromosome 3) are useful for identifying coding sequences and in the
 CC identification of biologically relevant splice junctions. The NHP
 CC polynucleotide sequences are useful in gene therapy, for detecting
 CC mutant NHPs or inappropriately expressed NHPs, for the diagnosis of

CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NHP.
 CC Complementary sequences of the NHP polynucleotides used in conjunction
 CC with PCR to screen libraries, isolate clones and prepare cloning and
 CC sequencing templates. Labeled NHP nucleotide probes can be used to screen
 CC a human genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests. The probe sequences also have use in defining and
 CC monitoring both drug action and toxicity. Oligonucleotides
 CC complementary to NHPs may encode or act as NHP antisense molecules, or
 CC may be used as part of ribozyme and/or triple helix sequences.
 CC The present sequence encodes an NHP.
 XX

SQ Sequence 2958 BP; 855 A; 694 C; 643 G; 764 T; 2 other;

Alignment Scores:
 Pred. No.: 1.33e-57 Length: 2958
 Score: 1384.00 Matches: 334
 Percent Similarity: 46.11% Conservative: 140
 Best Local Similarity: 32.49% Mismatches: 244
 Query Match: 20.96% Indels: 24
 DB: Gaps: 14

US-09-754-997A-2 (1-1252) x ABX71436 (1-2958)

Qy 9 GlyLeuLeuValLeuThrPheCysLeuSerAlaArgGlyGluLeuProLeuProGln 28
 Db 43 GGGATGCTGCTCGCGGCTCTCTGCTCTGCTGCTGCTCAGT-----CCTTTGCCAGGA 96
 Qy 29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValIleLeuGlyPro 47
 Db 97 GTGTGGTGTCTTAGCGAACTGTCTTTGTAAAGAACACACAGCATGTAACTGTACACAGA 156
 Qy 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProThrArg 67
 Db 157 AAGGACCCAGTCGTCTTTAGATTGCCAGGCTCACGGAGAGTT-----CCTATTAAAG 207
 Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro 87
 Db 208 GTACATGGTTGAATAATGAGCGCAAAATCTCTGMAAATAAACCGATCGAGGTCTTCTTCT 267
 Qy 88 AsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlnAspSerAspAspGluAla 107
 Db 268 AACGGCTCTTTATACATCAGT-----GAGGTGAAGGCGGAGGAGGAGCAGTCC 318
 Qy 108 LeuArgIleTrpLysValThrGluGlySerTyrosLeuAlaHisSerProLeuGly 127
 Db 319 -----GATGAAGGATTTTATCAGTGTCTGGCAATGAACAAATATGGA 360
 Qy 128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
 Db 361 GCCATTCTTAGTCAAAAGCTCATCTTGCCTTATCAACTATTCTGCAATTGAGTCCAG 420
 Qy 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
 Db 421 CCAATTTCCACTGAGGTCCACGAAGTGGAGTGTCTGATTTGCATTCAGATTTTCATCC 480
 Qy 168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
 Db 481 CACCCCTCTGCAGTCATAACATGAGGAGTTCAATCGGACAACTCTTACCTATGACTATGGAC 540
 Qy 187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
 Db 541 AGGATACTGCCCTTACCACAGAGGATTTGAGATCTATGATGTACCCAAAGGGATTTCT 600
 Qy 207 GlySerTyArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
 Db 601 GGAAATTCTGTTGTTGTTGCTGCCACTGTAGCCACCGACGTAAAGATGTAGAGGCTCG 660
 Qy 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValIleVal 246
 Db 661 CTAATCTGTGATT-----CCAGCTAAGGAGTCAAAATCCTTCCACACACACCAATATA 714

Qy	247	AlaAlaProGluAenThrThrValValSerGlyGlnAenValValMetGluCysValAla	266
Db	715	GCAGGTCACAGAACATAAACAACATCTCTTCATCAGACTGTAGTTTTGGAAATGCATGCC	774
Qy	267	SerAlaaspProThrProPheValSerTipValArgGlnAaspGlyIysProIleSerThr	286
Db	775	ACAGGAATCCCAACCAACATCATTTCTTGAGCGCCCTTCATCACAATCCATGTGATGC	834
Qy	287	--- AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis	305
Db	835	TTTAACTACTCGGTACTTGGAAATGTAATCATGTATATCTGATCTCAGGCTACCAACAT	894
Qy	306	SerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAaspPheAlaThrAlaAla	325
Db	895	GCTGGAGTATATGTTCTCGGGCCACTACCCCTGGCACACGCCAACTTTACAGTTGCTATG	954
Qy	326	AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg	345
Db	955	GCAACTTTAACTGTATTAGTCTCTCTTCATTTGTAATGGCCAGAAAGTTTAAACAAGG	1014
Qy	346	ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu	365
Db	1015	CCTCGACTGGCACTGCTCGATTGTGTGTGTCAGGCAGAGNAATCCCTCTCCCAAGATG	1074
Qy	366	HisTrpLeuHisAaspGlyIleProLeuArgProAenGlyArgValIysValGlnGlyGly	385
Db	1075	TCATGGTTGAAATGGAAGCAACATCAATCGAATCGTAGAATTAATAATGTAC-----	1128
Qy	386	GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAaspAlaGlyTyrTyrGlnCysVal	405
Db	1129	AACAGTAAATGGTAAATTAACCAAGATTATTCCTGAAGATGATGCTATTTATCATGTGCATG	1188
Qy	406	AlaGluAenSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValAlaArgGlu	425
Db	1189	GCTGAGAAATAGCCAAAGATCTATTTATCTAGAGCCAGACTGACTGTAGTGTGTGCGAA	1248
Qy	426	GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu	445
Db	1249	GACAGACCCAGTGTCTCCCTAATGTATGTCGTGAACCATGTCAAGCTCAGCCATCTCT	1308
Qy	446	ValAlaTrpGluArgProGluLeuHisSerGluClnIleIleGlyPheSerLeuHisTyr	465
Db	1309	TTAGCCTGGGAGGCCACCTTTAATTCAGACAAAGTCATTCCTATTCTGTACACTAC	1368
Qy	466	GlnLysAlaArgGlyValAaspAenValGluTyrGlnPheAlaValAenAenAaspThrThr	485
Db	1369	ATGAACAGCAGAGTTTAATATGAGAGATATCAAGTAGTCATCGGAATATGACACAACT	1428
Qy	486	GluLeuGlnValArgAaspLeuGluProAenThrAaspTyrGluPheTyrValAlaAlaTyr	505
Db	1429	CATTATATTATTGATGACTTAGAGCCCTGCCAGCAATTATCTTTCTACATTTGTAGCATAT	1488
Qy	506	SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAaspAaspVal	525
Db	1489	ATGCCAATGGGAGCCAGCCAGATGCTCGACCATGTGACACAGATACTCTAGAGGATGAC	1548
Qy	526	ProSerAlaAlaProGlnLeuThrLeuSerSerProAenProSerAaspIleArgValAla	545
Db	1549	CCCAGA-----	1554
Qy	546	TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGluTyr	565
Db	1554	-----	1554
Qy	566	GlyLeuGlyLysGluAaspGlnValPheSerThrGluValProGlyAenGluThrGlnLeu	585
Db	1554	-----	1554
Qy	586	ThrLeuAenSerLeuGlnProAenLysValTyrArgValArgIleSerAlaGlyThrGly	605
Db	1554	-----	1554
Qy	606	AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAenGln	625

1554	Db	----	----	1554
626	Qy	SerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuVal	645	
1554	Db	----	----	1554
646	Qy	ValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrpGly	665	
1554	Db	----	----	1554
666	Qy	GluValGlyThrGluGluGluAlaAspGlyAspArgProProGlyGlyArgGlyAspGln	685	
1554	Db	----	----	1554
686	Qy	AlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyrGluLeuThrGln	705	
1554	Db	----	----	1554
706	Qy	LeuValProGlyArgProTyrGluValLysValAlaPheAsnLysHisGluAspGly	725	
1555	Db	-----AGAAATATCATGTGAGACTCTCTGCTTACAAACAATAGACGATGGC	1602	
726	Qy	TyrAlaAlaValTrpLysGlyLysThrGlnLysAlaProThrProAspLeuProIleGln	745	
1603	Db	-----GATCAGACTCTCAGCACTCCAGGATCGGTCTCTTT	1647	
746	Qy	ArgGly-----ProProLeuProProAlaHisValHisAlaGluSerAsnSer	761	
1648	Db	CGTGATCGCATGTGTCCTCTCCACCACCACTCTCTATGCGAAGGCTTAACACC	1707	
762	Qy	SerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsn	781	
1708	Db	TCATCTCCATCTCTCGCACTGAGAGGCGCTGCATTCACCGCTGCACAATCATTAAC	1767	
782	Qy	TyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyrThr	801	
1768	Db	TACACCATCCGCTGTAATCTCTGTGGCTGCAGATGCTCTCTTGGTTCTGTACTCTCAA	1827	
802	Qy	SerSerGlyGluAspIleLeuIleGlyGlyLeuLysProPheThrLysTyrGluPheAla	821	
1828	Db	ACATCAGAAATCAATGTTGGTTCAAGGCTAGAACAAACAAATAACGAAATTTGCC	1887	
822	Qy	ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr	841	
1888	Db	GTTCGATTACATGTGGATCAGCTTTCAGTCTCTTGGAGCCCTGTAGTCTACCATTTACT	1947	
842	Qy	LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer	861	
1948	Db	CTTCCAGAGCACCAGCAGGCCACCACTGGAGTAAAGTGACATTAATAGAGGATGAC	2007	
862	Qy	ThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyrLeu	881	
2008	Db	ACTGCCCTGTTCTTGGAAACCCCTGTATGGCCCAAGAACAGTTGTGACCCGCTTACT	2067	
882	Qy	IleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeuThrThrGlu	901	
2068	Db	ATCTTATATGATCTAGGAAGCCCTGGATTGCAGGAGTGGCAGGCTTACACCGTGAA	2127	
902	Qy	GlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLys	921	
2128	Db	GGGGCAATAACCATGGCTTTGCTAGAAAACTTGTGTAGCAGGAAATATGTACATTTGTCAAG	2187	
922	Qy	MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGluAspValIleThr	941	
2188	Db	ATATCTGCATCCAAATGAGGTGGGAGAAGGACCCCTTTTCAAATTCCTGTGGAGCTGGCAGTA	2247	
942	Qy	LeuGlnGluThrPheSerAspSer-----	949	
2248	Db	CTTCCAAGGNAACCTCTGTAATCAAAATCAGAGGCCACCGTTTAGATTCTGTGTATGCC	2307	
950	Qy	-----LeuAspValHisAlaValThrGlyIleIleValGly	961	

Db	2308	AAAGTTTATTGAGGATATTACCATCTGGACCCAAAAATCAATGACTGGCAITGCTGTAGGT	2367
Qy	962	ValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGlnSerSer	981
Db	2368	GTGGCATAGCCTTGACCTGCATCCTCATCTGTGTCTTCATCTTGATATACCGAAGTAA	2427
Qy	982	HisArgGluAla-----LeuProGlyLeu	989
Db	2428	GCCAGGAATCATCTGCTTCCAGACGGCACAGAAATGGAACCTCAACAGTTACCTCGTACC	2487
Qy	990	SerSerSerGlyThrProGlyAsn	997
Db	2488	AGTGCCTCCTTAGCTAGTGGAAAT	2511

Search completed: October 14, 2003, 06:22:47
Job time : 680 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 14, 2003, 06:11:10 ; Search time 872 Seconds

(without alignments)
3725.578 Million cell updates/sec

Title: US-09-754-997A-2

Perfect score: 6604

Sequence: 1 MARADTGRGLLVLTCLLSA.....CPLTVSPSLPRAPVSSAQVP 1252

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1731049 seqs, 1297405648 residues

Total number of hits satisfying chosen parameters: 3462098

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09754997/runat_14102003_061101_9762/app_query.fasta_1.1415
-DB=Published Applications NA -QWFI=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODS=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09754997@cgn_1_1413 @runat_14102003_061101_9762
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6604	100.0	3756	10	US-09-754-997A-45 Sequence 45, Appl

2	6604	100.0	6176	10	US-09-754-997A-1	Sequence 1, Appl
3	5746	87.0	3753	12	US-10-239-663-29	Sequence 29, Appl
4	5610	84.9	3741	10	US-09-908-193-1	Sequence 1, Appl
5	4903	74.2	2796	10	US-09-754-997A-3	Sequence 3, Appl
6	3167.5	48.0	2223	12	US-10-239-663-28	Sequence 28, Appl
7	1747.5	26.5	3453	14	US-10-231-353-7	Sequence 1, Appl
8	1746.5	26.4	3210	14	US-10-231-353-1	Sequence 1, Appl
9	1746.5	26.4	3874	14	US-10-231-353-31	Sequence 31, Appl
10	1691.5	25.6	2976	14	US-10-231-353-11	Sequence 11, Appl
11	1691.5	25.6	3219	14	US-10-231-353-17	Sequence 17, Appl
12	1472	22.3	825	10	US-09-754-997A-5	Sequence 5, Appl
13	1384	21.0	2715	14	US-10-231-353-9	Sequence 9, Appl
14	1384	21.0	2958	14	US-10-231-353-3	Sequence 3, Appl
15	1329	20.1	2481	14	US-10-231-353-15	Sequence 15, Appl
16	1329	20.1	2724	14	US-10-231-353-19	Sequence 19, Appl
17	1144	17.3	2139	14	US-10-231-353-21	Sequence 21, Appl
18	1144	17.3	2382	14	US-10-231-353-27	Sequence 27, Appl
19	951.5	14.4	5297	12	US-10-007-926A-270	Sequence 270, App
20	909	13.8	1875	14	US-10-231-353-23	Sequence 23, Appl
21	781.5	11.8	1644	14	US-10-231-353-25	Sequence 25, Appl
22	781.5	11.8	1887	14	US-10-231-353-29	Sequence 29, Appl
23	647.5	9.8	1143	14	US-10-231-353-3	Sequence 3, Appl
24	624.5	9.5	7718	14	US-10-175-523-192	Sequence 192, App
25	607	9.2	8232	14	US-10-198-846-10976	Sequence 10976, A
26	592.5	9.0	909	14	US-10-231-353-13	Sequence 13, Appl
27	584.5	8.9	4277	11	US-09-946-374-57	Sequence 57, Appl
28	584.5	8.9	4277	12	US-10-015-387A-57	Sequence 57, Appl
29	584.5	8.9	4277	12	US-10-006-130A-57	Sequence 57, Appl
30	584.5	8.9	4277	12	US-10-199-672-439	Sequence 439, App
31	584.5	8.9	4277	12	US-10-006-172A-57	Sequence 57, Appl
32	584.5	8.9	4277	12	US-10-187-749-439	Sequence 439, App
33	584.5	8.9	4277	12	US-10-194-457-439	Sequence 439, App
34	584.5	8.9	4277	12	US-10-184-642-439	Sequence 439, App
35	584.5	8.9	4277	12	US-10-196-747-439	Sequence 439, App
36	584.5	8.9	4277	12	US-10-015-392A-57	Sequence 57, Appl
37	584.5	8.9	4277	12	US-10-017-253A-57	Sequence 57, Appl
38	584.5	8.9	4277	12	US-10-173-689-439	Sequence 439, App
39	584.5	8.9	4277	12	US-10-173-690-439	Sequence 439, App
40	584.5	8.9	4277	12	US-10-173-691-439	Sequence 439, App
41	584.5	8.9	4277	12	US-10-173-692-439	Sequence 439, App
42	584.5	8.9	4277	12	US-10-173-694-439	Sequence 439, App
43	584.5	8.9	4277	12	US-10-173-698-439	Sequence 439, App
44	584.5	8.9	4277	12	US-10-173-699-439	Sequence 439, App
45	584.5	8.9	4277	12	US-10-173-707-439	Sequence 439, App

ALIGNMENTS

RESULT 1

US-09-754-997A-45
; Sequence 45, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 3756
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-754-997A-45

Alignment Scores:

Pred. No.: 0 Length: 3756

Score:	6604.00	Matches:	1252
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0
US-09-754-997A-2 (1-1252) x US-09-754-997A-45 (1-3756)			
Qy	1	MetAlaArgAlaAspThrGlyArgGlyLeuValLeuValLeuThrPheCysLeuLeuSerAla	20
Db	1	ATGGCGGGGGGACACGGGCGGGGCTCTGGTGTGACCTTCTGGCTGCTCTGGCG	60
Qy	21	ArgGlyGluLeuProLeuProGlnGluThrThrVallysLeuSerCysAspGluGlyPro	40
Db	61	CGCGGGAGCTGCCATTGCCCCAGGACAACTGTCAAGCTGAGCTGTGATGAGGGACCC	120
Qy	41	LeuGlnValIleLeuGlyProGlnAlaValValLeuAspCysThrLeuGlyAlaThr	60
Db	121	CTGCAAGTGATCTCTGGGCTCTGAGCAGCTGTGGTGTGGACTGTGACTTTTGGGGGCTACA	180
Qy	61	AlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGluHis	80
Db	181	GCTGCTGGGCTCTCCGACCGGGTGACATGGAGCAAGGATGAGACACTGTACTAGAGCAT	240
Qy	81	GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu	100
Db	241	GAGAACTTGACCTGCTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAA	300
Qy	101	AspSerAspAspGluAlaLeuArgIleTrpLysValThrGluGlySerTrpSerCys	120
Db	301	GACAGCGATGATGAGGAAGCTCTTAGGATCTGGGAAGGTCACTGAGGGCAGCTATTTCCTGT	360
Qy	121	LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThr	140
Db	361	CTGGCCCAACGCCCTAGAGTGTGGCCAGCCAGGTGTGTGTGCTGACCTTGTGCCACA	420
Qy	141	LeuGluAspPheSerLeuHisProGluSerGlnIleValGluAsnGlyThrAlaArg	160
Db	421	CTCGAAGACTTCTCTCCACCCGAGTCCAGATTTGTGGAGGAGAACGGGACACACGC	480
Qy	161	PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnVal	180
Db	481	TTTGAATGCCACCAAGGGCTTCCAGCCCCCATCATTAATCTTGGGAAAGGACAGGTG	540
Qy	181	ThrValProGluLeuProArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAsp	200
Db	541	ACCGTGCTTGAGGAGCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCCATAG	600
Qy	201	ValGlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArg	220
Db	601	GTCCAGGACAGTGTGAGGCTCTACCGCTGGTGGCCACCAATTCAGCCCGCCCAACA	660
Qy	221	PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly	240
Db	661	TTTCAGCCAGGAGGCTCTGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAAGGGG	720
Qy	241	GlnAspValValIleValAlaProGluAsnThrThrValValSerGlyGlnAsnVal	260
Db	721	CAGGATGGTCAATTTGTGGCAGCCCGAGAAACACCGGTAGTGTCTGGACAGAAATGTA	780
Qy	261	ValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp	280
Db	781	GTGATGGAGTGGTGGCTCTGCTGACCCCACTTTTGTGTCTTGGGTCCGACAGGAT	840
Qy	281	GlyLysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSer	300
Db	841	CGAAAGCCTATCTCCACCGATGTATGCTGTCTGGCCCGGACCAATCTACTCATGCCAGC	900
Qy	301	AlaGlnProArgHisSerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAsp	320
Db	901	GGCAGACCTCGGACCTCTGGAGTCTATGTCTGCCGAGCCCAACAGCCCTCTCACGGTGC	960
Qy	321	PheAlaThrAlaAlaGluLeuArgValLeuAlaProAlaIleSerGlnAlaPro	340

Db	961	TTGCGCACCTGCGGCTGCTGAGCTCCGAGTCTTGTCTGCCCGCCAGCCATCTCGCAGGACCC	1020
Qy	341	GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu	360
Db	1021	GAGCGCTCTCGCGACCGCGGCGAGCACCGCGCTTCTGTGTCCGGGGCTTCGGGGAG	1080
Qy	361	ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgVal	380
Db	1081	CCAGGCCCGCGCTGCACCTGGCTGCACGCGGATCCGTTGGACCAATGGCGCGCTC	1140
Qy	381	LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly	400
Db	1141	AAAGTGCAGGGCGGTGGCGCAGCTTGGTTCATCATCAGATCGGCTGCGAGGACGTGGC	1200
Qy	401	TyrTrpGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAla	420
Db	1201	TACTACGATGCTAGCAGAAACAGCGCGGAACTGCTGTGCGCTGTGCGCCCTGGCG	1260
Qy	421	ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer	440
Db	1261	GTAGTGTGCGCGAGGGGCTGCCAGCGCCCGACTCGGCTCACAGCCACCGCGCTGAGC	1320
Qy	441	SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleGly	460
Db	1321	AGCTCTCTGTGCTGGTGGCTTGGAGCGGCTGAGTTGCACAGCAGCAAAATCATTTGGC	1380
Qy	461	PheSerLeuHisTrpGlnLysAlaArgGlyValAspAsnValGluTrpGlnPheAlaVal	480
Db	1381	TTCTCTTCTTACATACCAAGGCAAGGGAGTGACATGTGGAGTACCAAGTTTGCAGTA	1440
Qy	481	AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPhe	500
Db	1441	AACAATGACACACAGAGCTGCGAGTTGGGACCTTGGAAACCAACACGATTAAGTTTC	1500
Qy	501	TyrValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis	520
Db	1501	TACGTGTGGCTACTCTCCAGCTGGGGCCAGCCGAACTCTCAGCCAGCCCTGGTGAT	1560
Qy	521	ThrLeuAspAspValProSerAlaProGlnLeuThrLeuSerSerProAsnProSer	540
Db	1561	ACATGGACGATGTCCCGAGCGACCCCGAGCTTACCTTGTCCAGCCCAACCCCTCG	1620
Qy	541	AspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLys	560
Db	1621	GACATCAGGTGGCATGGCTGCCCTGCCCTCCAGCTGAGCAATGACAGGTGCTGAG	1680
Qy	561	TyrLysIleGluTrpGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly	580
Db	1681	TACAAGATAGATGACGGTTTGGGAGAGAGATCAGGTTTTTCTCCACCGAGGTGCTGGA	1740
Qy	581	AsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIle	600
Db	1741	AATCAGACACAACTTACGTTAAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGGAT	1800
Qy	601	SerAlaGlyThrGlyAlaGlyTrpGlyValProSerGlnTrpMetGlnHisArgThrPro	620
Db	1801	TCAGCTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGTGATGACAGCAGACACT	1860
Qy	621	GlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys	640
Db	1861	GGTGTGCACACACAGAGCCATGTTCCCTTTCGCCCTGCGAATTTGAAGGTGAGGCAAG	1920
Qy	641	MetGluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTyr	660
Db	1921	ATGAGTCCCTGGTGGTGTATGGACGCCGCCCTCTACCCCAACCCAGATCTCTGGATAC	1980
Qy	661	LysLeuTrpTrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProProGly	680
Db	1981	AAACTCTACTGGGAGAGGTGGGAACAGAGAGGAGGCGAGTGTGTACCGCCGCCACAGG	2040
Qy	681	GlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGln	700

Db 2041 GGTCTGGAGATCAAGCTTGGGACGTCGGGCGCTGGAGGAAAGTGAAGCAG 2100
Qy 701 TyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAen 720
Db 2101 TATGAATCAGCCAGTACCTGCTGGCAGGCGGTACAGGTGAAGCTCGTAGCTTCAAC 2160
Qy 721 LysHisGluAspGlyTyrAlaValTyrLysGlyLysThrGluLysAlaProThrPro 740
Db 2161 AAACAGAGAGCGGTACCTGCTGTGTGAAGGGCAAGCAGGAGAGCGGCCAGCCA 2220
Qy 741 AspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAen 760
Db 2221 GACCTGCTATCCAGAGGGGGCCACCGCTGCTCTGCCCCATGTCCACGACAGTCAAC 2280
Qy 761 SerSerThrSerIleTyrLeuArgTyrLysLysProAspPheThrThrValLysIleVal 780
Db 2281 AGCTCCACTTCCATTTGGGTTCGGTGAAGAGCCAGACTTTACCACTGTCAAGATTGTC 2340
Qy 781 AsnTyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThrTyr 800
Db 2341 AACTACACTGTACGCTTCGGCCCCCTGGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT 2400
Qy 801 ThrSerSerGlyGluAspIleLeuIleGlyGlyLeuLysProPheThrLysTyrGluPhe 820
Db 2401 ACCAGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACGAGTTT 2460
Qy 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840
Db 2461 GCGGTACAGTCCACGAGTGGATATGATGGGCTTTGGCTCCGTGTAGAACGCTCC 2520
Qy 841 ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
Db 2521 ACCCTGCTGACGGGCTTTCAACACCTCTCTGACCTCGGCTCGAGCCCCCTGCACCA 2580
Qy 861 SerThrValArgLeuHisTyrCysProThrGluProAsnGlyGluIleValGluTyr 880
Db 2581 TCACCGTTTGGTTACACTGGTGTCCCCCAGCGGCCAATGGTGGAGATTGTGGAGTAT 2640
Qy 881 LeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeuThrThr 900
Db 2641 CTAATTTCTCTACAGCAACCAACACACCCAGCCCGAACACCACTGGTCAACCACA 2700
Qy 901 GluGlyAsnIlePheSerIleGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
Db 2701 GAGGGAAACATCTTCAGTGCAGAGGTCCATGGCTAGAGAGTCACTCGGTATTTCTTC 2760
Qy 921 LysMetGlyValArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle 940
Db 2761 AAGATGGAGCCCGCACAGGTGGGCGCTGGGCGCTTTCCGCTTGCAGGATGTGATT 2820
Qy 941 ThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIleIleVal 960
Db 2821 ACTCTGCAAGAGACATTTCTCAGACTCTTGGATGTGCACGCGTCAOGGCGCATCTG 2880
Qy 961 GlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
Db 2881 GGTGTCTGCTGGGCTTCTCTGCTCTCTGGGCTGTCATGTGTGCTGGCTTACGACAAAGC 2940
Qy 981 SerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeu 1000
Db 2941 TCCACAGGGAAGCCCTTCCCGATTGTCTCTCAGGCGACCCCGAGAAACCCAGCGCTC 3000
Qy 1001 TyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeu 1020
Db 3001 TACACAGAGCTCGGCTTGGGCTCCAGTGTCTCTGCTGCTCCCATGAGTTGAGTCCCTC 3060
Qy 1021 ValHisProArgProGlnAspTyrSerProProProSerAspValGluAspLysAlaGlu 1040
Db 3061 GTGCTATCTCTGCTCCCGAGATTGGTCCCCACCACTCAGATGTGCAAGCAAGGCTGAA 3120
Qy 1041 ValHisSerLeuMetClyGlySerValSerAspCysArgGlyHisSerLysArgLysIle 1060
Db 3121 GTACACAGCTTTATGGTGGCGAGTGTTCAGATTGCCGGGGCCACTTCCAAAGAGAAGATC 3180

Qy 1061 SerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGlnLeuPro 1080
Db 3181 TCCTGGGCTCAGGCAGGGGGACCAAACTGGGCGAGGCTCTCTGGCAGGCTGTGAGCTGCC 3240
Qy 1081 GlnGlySerGlyProAspProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
Db 3241 CAGGATAGTGGTCCAAAGCCCGGCTCTGACCCGTGTCTGTGCTCTCCAGCGGAAACCGGG 3300
Qy 1101 GlnThrLeuLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLysLys 1120
Db 3301 CAGACACTGCTGCTGCAAGCCCTGGTGTATGACGGCATAAAGAGCAACGGGAGAAAGAG 3360
Qy 1121 ProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGly 1140
Db 3361 CCGTCCCCAGCTGCGAGGAATCAGGTGGAGCTGAGGTCAATTGTCTCACTCCGACTTCGGT 3420
Qy 1141 AlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThr 1160
Db 3421 GCATCCAAAGGATGTCTGACCTCCACCTCAAGACCTGGAGCCAGAGAAACCACTGACT 3480
Qy 1161 AlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyValAspTyrLeu 1180
Db 3481 GCAGAGACTCTGCCCTTCCAGCTCTGGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTG 3540
Qy 1181 GlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
Db 3541 GCGAGGAGCTGGGAGGGTGCACCAACCAACCAAGTGGGCGCCAGAGGCTCACCTGCTTG 3600
Qy 1201 ProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAlaIleGlu 1220
Db 3601 CCAGAAGCAGCAGCTGCTCTGCTCTGCTCAGACCTCCAGCCCGCAGCAGCTGTATAGAG 3660
Qy 1221 GluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSer 1240
Db 3661 GAGGCGCTTGGAAAGTGTGCGCCCAAGCCCTGTGTCTCTTAACTGAGTCCGCCCAAGC 3720
Qy 1241 LeuProArgAlaProValSerSerAlaGlnValPro 1252
Db 3721 CTTCCAGGCGCCCTGTCTCTCTGCTCAGGTCCCC 3756

RESULT 2

US-09-754-997A-1
; Sequence 1, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6176
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(3756)
US-09-754-997A-1
Alignment Scores:
Pred. No.: 0 Length: 6176
Score: 6604.00 Matches: 1252
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

DB:	10	Gaps:	0
US-09-754-997A-2 (1-1252) x US-09-754-997A-1 (1-6176)			
Qy	1	MetAlaArgAlaAspThrGlyArgGlyLeuLeuValLeuThrPheCysLeuLeuSerAla	20
Db	1	ATGCGCGGCGGACACGGCGCGGCTCTGTGTGTGACCTTCTCCCTGTGTCTCGCG	60
Qy	21	ArgGlyGluLeuProLeuProGlnGluThrThrVallyLeuSerCysAspGluGlyPro	40
Db	61	CGCGGGAGAGTGCCTATTTGCCCCAGAGACAACTGTCAAGCTGTGACTGTGATGAGGGACCC	120
Qy	41	LeuGlnValIleLeuGlyProGluGlnAlaValIleValLeuAspCysThrLeuGlyAlaThr	60
Db	121	CTGCAAGTGATCTCTGGGCCCTGACAGCGCTGTGTGTGTGACTGTCACTTTGGGGGCTACA	180
Qy	61	AlaAlaGlyProProThrArgValThrTrpSerIysAspGlyAspThrValLeuGluHis	80
Db	181	GCTGCTGGGCTCCGACAGCGGTGACATGGAGCAAGGATGGAGACACTGTACTAGAGCAT	240
Qy	81	GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu	100
Db	241	GAGAACCTGCACCTGCTACCCAAATGGCTCCCTGTGGCTGTCCCTCACCCCTAGAGCAAGAA	300
Qy	101	AspSerAspAspGluGluAlaLeuArgIleTrpIysValThrGluGlySerTrpSerCys	120
Db	301	GACAGCGATGATGAGGAGCTCTTAGGATCTGGAGGTCCTAGGAGGCGAGCTATTCTCTGT	360
Qy	121	LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValValValValAlaThr	140
Db	361	CTGGCCACACAGCCGCTAGGAGTGTGTGGCCAGCCAGCTGTGTGTGTCAAGCTTGGCCACA	420
Qy	141	LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg	160
Db	421	CTCGAAGACTTCTCTGTGCACCCCGAGTCCAGATTTGTGGAGGAGAACGGGACAGCAGC	480
Qy	161	PheGluCysHisThrIysGlyLeuProAlaProIleIleThrTrpGluIysAspGlnVal	180
Db	481	TTTGAAATGCCACACCAAGGCGCTTCCAGCCGCCATCATATTCTGGGAAAGGACACAGGTG	540
Qy	181	ThrValProGluGluProArgLeuIleThrLeuProIysTrpLeuLeuGlnIleLeuAsp	200
Db	541	ACCGTGCCTGAGGAGCCCGGCTCATCACTCTCCCAAGTGGCTCTCCAGATCCTAGAT	600
Qy	201	ValGlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArg	220
Db	601	GTCCAGGACAGTGTATGACGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCACGA	660
Qy	221	PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly	240
Db	661	TTTCCGCCAGGAGGCTCTGCTCACTGTGGCCCTCAGAGGTCCTTTGGAGGCTTACCAGGGGG	720
Qy	241	GlnAspValValIleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnVal	260
Db	721	CAGGATGTGGTCAATTTGTGGCAGCCCGAGAGAACCCAGCGTAGTGTCTGGACAGAAATGTA	780
Qy	261	ValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp	280
Db	781	GTGATGGAGTGGTGGCTCTGCTGACCCCAAGCTTTGTGTCTCTGGGTCCGACAGGAT	840
Qy	281	GlyIysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSer	300
Db	841	GGAAAGCCTATCTCCAGGATGTATGCTCTGGGCCGGACCAATCTACTACTCATGCCGACG	900
Qy	301	AlaGlnProArgHisSerGlyValTrpValCysArgAlaAsnIysProLeuThrArgAsp	320
Db	901	GGCAGCCTCGGCACCTCTGGAGTCTATGCTCTGGCCGAGCAACAGCCCTCACCGGTGAC	960
Qy	321	PheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro	340
Db	961	TTTCGCCACTGGCGGTGCTGAGTCCGAGTGTGTGTCGCCCGAGCCATCTCTCGCAGGACCC	1020
Qy	341	GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu	360
Db	1021	GAGCGCTCTCGCGACCGCGGCGCAGACCGCGCGCTTGTGTGTGCGCGCGCTCCGCGGAG	1080
Qy	361	ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgVal	380
Db	1081	CCACGCGCCCGCTGCACCTGGCTGCGACGACGGGATCCCGTTGGACCCCAATGGCGCGCTC	1140
Qy	381	LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly	400
Db	1141	AGGTGTCAGGCGGTGGCGCAGCTTGGTCATCACTCAGATCGGCTTCGAGGACGCTGGC	1200
Qy	401	TrpTrpGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAla	420
Db	1201	TACTACAGTGGCTAGCAGAAAACAGCGGGGAATGCTGTGTGCGCTGCGCCCTGGCG	1260
Qy	421	ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer	440
Db	1261	GTAGTGGTGGCGAGGGGCTGCCAGCGCCCGCTGAGTTGCACAGCGCAAAATCATTTGGC	1320
Qy	441	SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGly	460
Db	1321	AGCTCTCTGTGTGTGGCTGGGAGCGGCTGAGTTGCACAGCGCAAAATCATTTGGC	1380
Qy	461	PheSerLeuHisTrpGlnIysAlaArgGlyValAspAsnValGluTrpGlnPheAlaVal	480
Db	1381	TTCTCTCTTCACTATCAAAAAGGCAAGGGAGTGGACAATGTGTGAGTACCAGTTTGCAGTA	1440
Qy	481	AsnAspAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPhe	500
Db	1441	AACAATGACACCAACAGAGCTGCAGGTTCGGGACCTGGAAACCCCAACGGAATATAGTTTC	1500
Qy	501	TyrValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis	520
Db	1501	TAGTGGTGGCTACTCTCCAGCTGGGGGCGCAGCGGAACCTCCAGCCCGAGCCCTGGTGCAT	1560
Qy	521	ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer	540
Db	1561	ACACTGGACGATGTCCCGAGCGCAGCACCCAGCTACCTTGTTCAGCCCGCAACCCCTCG	1620
Qy	541	AspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuIys	560
Db	1621	GACATCAGGCTGGCATGGCTGCCCCCTCCAGCTTGAGCAATGACAGGTGTCTGAAG	1680
Qy	561	TyrIysIleGlyTrpGlyLeuGlyIysGluAspGlnValPheSerThrGluValProGly	580
Db	1681	TACAGATAGTACGGTTTGGGAGGAGAGATCAGGTTTTCTCCACCGAGGTGCTTGA	1740
Qy	581	AsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnIysValTrpArgValArgIle	600
Db	1741	AATGAGACACACTTACGTTAAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGAT	1800
Qy	601	SerAlaGlyThrGlyAlaGlyTrpGlyValProSerGlnTrpMetGlnHisArgThrPro	620
Db	1801	TCAGCTGGCATCTGGCTGTGTATGGAGTCCCTTCTCAGTGGATGCGACAGAGACCT	1860
Qy	621	GlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuIysValArgAlaIys	640
Db	1861	GGTGTGCACAAACAGGACCATGTTCCCTTGGCCCTCGAGAAATGAGGTGAGGCGCAAG	1920
Qy	641	MetGluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTrp	660
Db	1921	ATGGAGTCCCTGGTGGTGTCTATGGCAGCCCGCCCTCACCCACCCAGATCTCTGGATAC	1980
Qy	661	LysLeuTrpTrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProProGly	680
Db	1981	AAACTCTACTGGGGAGAGGTGGGAACAGAGAGGAGGAGATGTTGTGACCGCCCGCCAGGG	2040
Qy	681	GlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuIysValIysValIysGln	700
Db	2041	GCTGTGGAGATCAAGCTTGGAGCTCGGGCCGCTGCGCTGGAAGAGAAAGTAGTAGCAG	2100
Qy	701	TyrGluLeuThrGlnLeuValProGlyArgProTrpTrpGluValIysLeuValAlaPheAsn	720

Db 2101 TATGAACGACCGCTAGTCCCTGGCAGGCGCTACGAGGTGAAGCTCGTAGCTTTCAC 2160
QY 721 LysHisGluAspGlyTyrAlaValTrpLysGlyLysThrGluLysAlaProThrPro 740
Db 2161 AAACACGAGGACGGCTACGCTGCTGTGTGAAGGCAAGACGAGAGGCGCCACGCGCA 2220
QY 741 AspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsn 760
Db 2221 GACCTCCCTATCCAGAGGGGGCCACCGCTGCCCTCCCTGCCATGTCCACGCGAGTCAAA 2280
QY 761 SerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleVal 780
Db 2281 AGCTCCACTTCCATTTGGCTTCGGTGGAGAGACCGAGCTTTACCACTGTGCAGATTGTC 2340
QY 781 AsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyr 800
Db 2341 AACTACACTGTACGCTTCGGCCCTGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT 2400
QY 801 ThrSerSerGlyGluAspIleLeuIleGlyLysLysProPheThrLysTyrGluPhe 820
Db 2401 ACCAGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACGAGTTT 2460
QY 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840
Db 2461 GCGGTACGTTCCACGAGTGGATATGATGGGCGCTTTGGCTCCGTGTAGAACGCTCC 2520
QY 841 ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
Db 2521 ACCCTGCTGACCGGCTTCAACACACTCTCTCTGACCTGGCCTGAGCCCTTGACACCA 2580
QY 861 SerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyr 880
Db 2581 TCCACCGTTTGGTTACACTGGTGTCCCGCCACGAGGCCCAATGGTGGAGATTGGAGTAT 2640
QY 881 LeuIleLeuTyrSerAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThr 900
Db 2641 CTAAATCTCTACAGCAACCAACCAACCCAGCCCGAACCAGTGGACACTGTCTCACCA 2700
QY 901 GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
Db 2701 GAGGGAACATCTTCAGTGGAGAGGTCCATGGCTAGAGGTGACACTCGGTATTTCCTC 2760
QY 921 LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle 940
Db 2761 AAGATGGAGCCGCGCACAGAGTGGGGCTGGGCGCTTTTCCGCTTGCAGGATGTGATT 2820
QY 941 ThrLeuGlnThrPheSerAspSerLeuAspValHisAlaValThrGlyIleIleVal 960
Db 2821 ACTCTGCAAGAGACATTTCTCAGACTCTTGGATGTGCAGCGCTCAGCGGCATCATCGTG 2880
QY 961 GlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
Db 2881 GGTGTCTGCTGGGCTTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
QY 981 SerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeu 1000
Db 2941 TCCACAGGGAAGCCCTTCCGCGATTGTCTCTCAGGACACCCAGGAAACCCAGCGCTC 3000
QY 1001 TyrThrArgAlaArgLeuGlyProProSerValProAlaHisGluLeuGluSerLeu 1020
Db 3001 TACACAGAGCTCGGCTTGGGCTTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
QY 1021 ValHisProArgProGlnAspTrpSerProProSerAspValGluAspLysAlaGlu 1040
Db 3061 GTGCATCTCTGCTCCAGGATTGGTCCCAACCACTTCAGATGTGAAGCAAGGCTGAA 3120
QY 1041 ValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIle 1060
Db 3121 GTACACAGCTTATGGTGGCAGTGTTCAGATTGGCGGGCCACTCCAGAGCAAGATC 3180
QY 1061 SerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuPro 1080
Db 3181 TCCTGGGCTCAGGACGAGGGGACCAAACTGGGAGGCTCTCTGGGAGGCTGTGAGCTGCC 3240

QY 1081 GlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
Db 3241 CAGGGTAGTGGTCCAAGCGCGCTCTGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300
QY 1101 GlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLysIle 1120
Db 3301 CAGACACTGCTCTGCAAGCCCTGGTGTATGACGGCATAAAGAGCAACGGGAGAAAG 3360
QY 1121 ProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGly 1140
Db 3361 CGGTCCCGACCTGCGAGGATCAGGTGGAGCTGAGGTCACTTGTCCACTCCGACTTGGT 3420
QY 1141 AlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThr 1160
Db 3421 GCATCAAAAGATGTCTGACCTCCACCTCCAAAGACCTGGAGCCAGAGGACCACTGACT 3480
QY 1161 AlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeu 1180
Db 3481 CAGAGACTCTGCTCTCCACGCTGGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTG 3540
QY 1181 GlyArgGluLeuGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
Db 3541 GGCAGGAGCTGGAGGGTGGCCACCAACACCAACAGTGGGCGGAGAGGCTCACCTGCTG 3600
QY 1201 ProGluAlaAlaSerAlaSerCysSerCysAspLeuGlnProSerThrAlaIleGlu 1220
Db 3601 CCAGAGCAGCAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3660
QY 1221 GluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSer 1240
Db 3661 GAGGCGCTTGGGAAAGCTGCGAGCCCAAGCCCTGTGCTCTTAACAGTCAAGCCCAAGC 3720
QY 1241 LeuProArgAlaProValSerSerAlaGlnValPro 1252
Db 3721 CTTCCAGGCGCTGCTGCTCTGCTGAGTCCCC 3756

RESULT 3

US-10-239-663-29
; Sequence 29, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabinick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 3753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-239-663-29

Alignment Scores:
Pred. No.: 0 Length: 3753
Score: 5746.00 Matches: 1098
Percent Similarity: 92.09% Conservative: 54
Best Local Similarity: 87.77% Mismatches: 96

Db 2098 AGACATATGAGCTGACCCAGCTAGTCTCGCCGCTGTACGAGGTGAAGCTCGTGGCT 2157
Qy PheAsnLysHisGluAspGlyTyrAlaAlaValTyrLysGlyLysThrGluLysAlaPro 738
Db 2158 TTCAACAAACATGAGATGGCTATGACGATGCGATGTGGAGGGCAGACGAGAGAGGCCCG 2217
Qy ThrProAspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGlu 758
Db 2218 GCACACAGATGCTATACAGAGGGGACCAACCCCTGCTCCAGCCCAAGTCCATCGGAA 2277
Qy SerAsnSerSerThrSerIleTyrLeuArgTyrLysProAspPheThrThrValLys 778
Db 2278 TCAACACAGCTCCACATCCATCTGGCTTCGGTGGTGGAAAGCCAGATTTCCACAGTCAAG 2337
Qy IleValAsnTyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThr 798
Db 2338 ATTGTCACTACACTGTGGCTTCAGCCCTCGGGGCTCAGGATGCCCTCCCTGGTCAAC 2397
Qy TyrTyrThrSerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrLysTyr 818
Db 2398 TATTACACAGTTCCTGGAGAGACATCCTCATTTGGCGGCTTGAAGCATTCCACCAATAC 2457
Qy GluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGlu 838
Db 2458 GAGTTTGCAGTGCAGTCTCACGGCGTGGACATGGATGGGCTTTTCGGCTCTGTGGTGGAG 2517
Qy ArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeu 858
Db 2518 CGCTCAACCTCGCTGACCGGCTCCACACCCCTCCAGCTCGACTGAGCCGCCCTG 2577
Qy ThrProSerThrValArgLeuHisTyrCysProProThrGluProAsnGlyGluIleVal 878
Db 2578 ACACCGTCCAGCGTTTCGGCTGCATCTGTGCCCCCACCACAGAGCCCAAGCGGAGATCGT 2637
Qy GluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeu 898
Db 2638 GAGTATCTGATCTGTATACAGCAGCAACACACAGCAGCTGAGCACCAAGTGGACCTTGTCTC 2697
Qy ThrThrGluGluAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyr 918
Db 2698 ACCACGAGGAAACATCTTCAGTGTCTGAGGTCCATGGCTCGAGAGCGACATCTCGTATC 2757
Qy PhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAsp 938
Db 2758 TTCTTCAAGATGGGGCGGCACAGAGTGGGACCTGGGCTTCTCCCGCTCGCAGAT 2817
Qy ValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIle 958
Db 2818 GTGATCACGCTCCAGGAGAGCTGTACAGACTCGCTGACATGCATCAGTCACGGGCATC 2877
Qy IleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArg 978
Db 2878 ATCGTGGGTGTGCTGGGCTCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2937
Qy GlnSerSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnPro 998
Db 2938 CGCAGCCCCCAGGAGATCTCTCCAGGCTGTCTCTCCAGCCACCCCGGGAATCCC 2997
Qy AlaLeuTyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGlu 1018
Db 2998 GCGCTGTACTCCAGAGCTCGCTTGGCCCCCCCCCAGCCCCCAGCTGCCATGAATTGGAG 3057
Qy SerLeuValHisProArgProGlnAspTyrSerProProProSerAspValGluAspLys 1038
Db 3058 TCCCTTGTGCACCCCATCCAGGACTGTGTCGCCGCCACCTCCAGACGCTGGAGGACAGG 3117
Qy AlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArg 1058
Db 3118 GCTGAGTGCACAGCCTTATGGTGGCGGTGTTCGAGGGCGGAGTCACTCCAAAAGA 3177
Qy LysIleSerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGlu 1078

Db 3178 AAGATCTCTCTGGGCTCAACCAAGCGGGCTGAGCTGGGTGGTCTCTGGGAGGCTGTGAG 3237
Qy LeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGly 1098
Db 3238 CTGCCCCCAG--GCAGGCCCCCGCGGCTCTGACCCCGGGCTCTGTCCTCTCTGCTGA 3294
Qy ThrGlyGlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArg 1118
Db 3295 ACTGGGCGAGACGCTCTGCTGCGAGGCTCTGGTGTACGACGCCATTAAGGGCAATGGAGG 3354
Qy LysLysProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAsp 1138
Db 3355 AAGAAGTCAACCCCGAGCTTCAGGAAACAGGTGGAGGCTGAAGTCAATTGTCCACTCTGAC 3414
Qy PheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluPro 1158
Db 3415 TTTAGTGCATCTAACCGGNAACCTGACCTCCATCTCCAGACCTCGAGCTGAGGCTGAGGACCCC 3474
Qy LeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAsp 1178
Db 3475 CTGCTCCAGAGGCTCTCATCTCATCTCGGGTGTGGGATCCAGGGCAGGGGCGAGCC 3534
Qy TrpLeuGlyArgGluLeuGlyCysGlnProThrThrSerGlyProGluArgLeuThr 1198
Db 3535 TGGCTGGACAGGAGTGGAGGGTGTGAGCTGGCAGCCCCCGGCGCCAGACAGACTTACC 3594
Qy CysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAla 1218
Db 3595 TGCTTCCAGAGGAGCCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3654
Qy IleGluGluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSer 1238
Db 3655 CTAGAGGAGACCCCTGGAGATAGTGCAGCTCAAAATCCCTCTCTCTCTCTCTCTCTCTCT 3714
Qy ProSerLeuProArgAlaProValSerSerAla 1249
Db 3715 CCAGCGTCCCGAGATCCCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3747

RESULT 4

US-09-908-193-1
; Sequence 1, Application US/09908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMKETS, RICHARD A.
; APPLICANT: ZERHUSEN, BRYAN
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3741
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: NOVX nucleic acid
 ; OTHER INFORMATION: with homology to NOPG-like protein
 ; NAME/KEY: modified_base
 ; LOCATION: (375)
 ; OTHER INFORMATION: a, c, t, g, other or unknown
 US-09-908-193-1

Alignment Scores:

Pred. No.: 0 Length: 3741
 Score: 5610.00 Matches: 1086
 Percent Similarity: 91.07% Conservative: 56
 Best Local Similarity: 86.60% Mismatches: 101
 Query Match: 84.95% Indels: 13
 DB: 10 Gaps: 8

US-09-754-997a-2 (1-1252) x US-09-908-193-1 (1-3741)

Qy 1 MetAlaArgAlaAspThr-----GlyArgGlyLeuLeuValLeuThrPheCysLeuLeu 18
 Db |||||
 Qy 1 ATGGCGGGGGGACGCGCGCGCGCGCGCGCTCTCGCGTTGACCTTCTGCTGCTGTG 60
 Db |||||
 Qy 19 SerAlaArgGlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCysAspGlu 38
 Db |||||
 Qy 61 GCCCGCGCGGGAGCTGCTGTTGCCCGAGAGACGACTGTGGAGCTGAGCTGGAGTG 120
 Db |||||
 Qy 39 GlyProLeuGlnValLeuLeuGlyProGlnGlnAlaValValLeuAspCysThrLeuGly 58
 Db |||||
 Qy 121 GGGGCACCTGCAGTGATCTGGGCCAGAGCGGCTGCAGTGTCTAACTGTAGCTGGGG 180
 Db |||||
 Qy 59 AlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeu 78
 Db |||||
 Qy 181 GCTGCTGCGCTGACCCCCACCGAGGTGACCTGGAGCAAGGATGGGACACCCCTGCTG 240
 Db |||||
 Qy 79 GluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuThrProLeuSerProLeuGlu 98
 Db |||||
 Qy 241 GAGCACGACACTTACACCTGCTGGCCAAATGGTTCCCTGTGGCTGTCCAGCCACTAGCA 300
 Db |||||
 Qy 99 GlnGluAspSerAspGluAlaLeuAlaLeuArgIleTrpLysValThrGluGlySerTyr 118
 Db |||||
 Qy 301 CCCAATGGCAGTGACGAGT--CAGTCCCTGAGGCTGTGGGG-GTCAATTGAGGCACTAT 357
 Db |||||
 Qy 119 SerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeu 138
 Db |||||
 Qy 358 TCGTGCTTAGCCACCGNCCCTCGAGTGTGTGGCCAGCCAGAGCTGCTGCTCAAGCTT 417
 Db |||||
 Qy 139 AlaThrLeuGluAspPheSerLeuHisProGlnSerGlnIleValGluGluAsnGlyThr 158
 Db |||||
 Qy 418 GCCAGTCTCGCAGACTTCTCTCGACCCGGAGTCTCAGACGGTGGAGGAGAACGGGACA 477
 Db |||||
 Qy 159 AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAsp 178
 Db |||||
 Qy 478 GCTCGCTTTGATGTCACATTTGAAGGGCTGCAGCTCCCATCTATTTCTGGAGAGAGAC 537
 Db |||||
 Qy 179 GlnValThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeuGlnIle 198
 Db |||||
 Qy 538 CAGGTGACATTGCTGAGGACCTCGGCTCATCGTCTTCCCAACCGCGCTCTTCAGATC 597
 Db |||||
 Qy 199 LeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAlaArg 218
 Db |||||
 Qy 598 CTGATGTTTACGAGAGTGTATGACAGGCCCTTACCGCTCGTGGGCCACCAACTCAGCTCGC 657
 Db |||||
 Qy 219 GlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThr 238
 Db |||||
 Qy 658 CAGCACTTACGACAGGAGGCCCTTACTGATGTGTGGCCACAGAGGTTTCTTGGCGTCCACC 717
 Db |||||
 Qy 239 ArgGlyGlnAspValValLeuAlaAlaProGluAsnThrThrValValSerGlyGln 258
 Db |||||
 Qy 718 AGGGGCGCAGGACGTGGTCAATTGTGGCAGCCCCCAGAGAACCAACAGTGTGTCTGGCCAG 777
 Db |||||
 Qy 259 AsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArg 278
 Db |||||
 Qy 778 AGTGTGTGTGATGGAATGTGTGGCCTCAGCTGACCCCAACCCCTTTTGTGTCTGTGGGTCCGA 837
 Db |||||

Qy 279 GlnAspGlyLysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuLeu 298
 Db |||||
 Qy 838 ---GACGGGAGGCCATCTCCACAGATGTATCTCTGGCGCCACCAACCTACTTAATT 894
 Db |||||
 Qy 299 AlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysProLeuThr 318
 Db |||||
 Qy 895 GCCAAGCGCGACGCTGGCAGCTCCGCGCTCTATGTCTGCCGCGCCCAACAGCCCGCAGC 954
 Db |||||
 Qy 319 ArgAspPheAlaThrAlaAlaGluLeuArgVal---LeuAlaAlaProIleLysSer 337
 Db |||||
 Qy 955 CGGACTTCCCACTGCAGCGCTGAGCTCCGTGTGTGTCTAGCGGCTCCCGCATCACT 1014
 Db |||||
 Qy 338 GlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAla 357
 Db |||||
 Qy 1015 CAGCGCGCCGAGGCGCTGTCGCGACGCGCGCGAGCACAGCGGCTTCTGTGTGCGCGCG 1074
 Db |||||
 Qy 358 SerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsn 377
 Db |||||
 Qy 1075 TCGGGGAGCGCGCGCGCGCTGCGCTGGCTGCACAAACGGGGCGCGCTGCGGCCCAAC 1134
 Db |||||
 Qy 378 GlyArgValLysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGln 397
 Db |||||
 Qy 1135 GGGCGGTCAAGGTCCAGGGCGCGGTGGCGGCTGTGTATCATCACAGATCGGCTGCAG 1194
 Db |||||
 Qy 398 AspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAla 417
 Db |||||
 Qy 1195 GACCGCGCTACTACAGTGGCTGGCTGAGAACAGCGCGGGAATGGCGTGGCTGCGCG 1254
 Db |||||
 Qy 418 ProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThr 437
 Db |||||
 Qy 1255 TCGCTGGCGCTGGTGGTGGCGAGGGGCTGCCAGCGCCCCACCGGGTCACTGTCTACG 1314
 Db |||||
 Qy 438 ProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGln 457
 Db |||||
 Qy 1315 CCATGAGCAGCTCCGCTGTGTGGTGGCTGGAGCGGCCGAGATGCAAGCGAGCAG 1374
 Db |||||
 Qy 458 IleIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyrGln 477
 Db |||||
 Qy 1375 ATCATCGGCTTCTCTCTCCACTACAGAGGACAGCGGGTATGACAAATGTGGAAATACCAG 1434
 Db |||||
 Qy 478 PheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAsp 497
 Db |||||
 Qy 1435 TTTTGCAGTGAACACACACACACAGAACTACAGGTTCCGGACCTCGGAACCCCAACACAGAT 1494
 Db |||||
 Qy 498 TyrGluPheTyrValValAlaTyrSerGlnLeuGlyValaSerArgThrSerSerProAla 517
 Db |||||
 Qy 1495 TATGAGTTCCTACGTGTGGCTTACTCCAGCTGGAGCCAGCGCCACTCCACCCCGCA 1554
 Db |||||
 Qy 518 LeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerPro 537
 Db |||||
 Qy 1555 CTGTGTCACACACTGGATGATGTCCCAGTGCAGCACCCCGCTCTCCCTGTCCAGGCCCC 1614
 Db |||||
 Qy 538 AsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGln 557
 Db |||||
 Qy 1615 AACCTTCCGACATCAGGGTGGGTGGCTGCCCTCGCCCCCAGCGCTGAGCAATGGGCGAG 1674
 Db |||||
 Qy 558 ValLeuLysTyrIleGlyLeuGlyValGluAspGlnValPheSerThrGlu 577
 Db |||||
 Qy 1675 GTGGTGAAGTACAGATAGAAATACGGTTTGGGAAAGGAATCAGATTTTCTTACTAGAG 1734
 Db |||||
 Qy 578 ValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArg 597
 Db |||||
 Qy 1735 GTGCGAGGAATGAGACACAGCTTATGCTGAACCTCGCTTCCAGCCCAACAGGTGTATCGA 1794
 Db |||||
 Qy 598 ValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHis 617
 Db |||||
 Qy 1795 GTACGGAATTCGGCTGTGTACAGCAGCGGCTTCCGGGCCCTCCCTCCAGTGTGATCATC 1854
 Db |||||
 Qy 618 ArgThrProGlyValHisGlnSerHisValProPheAlaProAlaGluLeuLysVal 637
 Db |||||
 Qy 1855 AGACGCCCAAGTATGACACCAACAGAGCCATGTCCCTTTTGGCCCTCGCAGAGTTGAAGGTG 1914
 Db |||||
 Qy 638 ArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGlnIle 657
 Db |||||

1915 CAGGCAAGATGGAGTCCCTGGTGTCTATGGCAGCCACCCCTCACCCACCCAGATC 1974
658 SerGlyTyrLeuTyrTrpGlyValGlyThrGluGluGluAlaAspGlyAspArg 677
1975 TCTGGCTACAACTATATGGCGGAGGTGGGGCTGAGGAGGAGGCAATGGGATCGC 2034
678 ProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLeuLys 697
2035 CTGCCAGGGGGCCGTGGAGACACAGGCTTGGGATGGGGCTGTCTCCGGCTCAAGAGAAA 2094
698 ValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuVal 717
2095 GTGAAGCAGTATGAGCTAGCCAGCTAGTCCCTGGCCGCTGTACGAGGTGAAGCTCGT 2154
718 AlaPheAenLysHisGluAspGlyTyrAlaAlaValTrpLysGlyThrGluLysAla 737
2155 GCTTTCAACAAACATGAGATGGCTATGACAGAGTGTGGAGGCGACAGACGGAGGGCG 2214
738 ProThrProAspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAla 757
2215 CCGGCACACAGATGCTATCCAGAGGGGACACCCCTGCTCCAGCCACGTCATGGC 2274
758 GluSerAenSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrVal 777
2275 GAATCAACAGCTCCACATCATCTGGCTTGGGTGGAAAGCCAGATTTCCACACAGTC 2334
778 LysIleValAenTyrThrValArgPheGlyProTrpGlyLeuArgAenAlaSerLeuVal 797
2335 AAGATTGTCAACTACATCTGGCTTCCAGCCCTGGGGGCTCAGGAATGCTCTCTGGTC 2394
798 ThrTyrTyrThrSerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrLys 817
2395 ACCTATTATAC---AGTTCTGGAGAGACATCTCTATTGGCGGCTTGAAGCCATTCACCAAA 2451
818 TyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValVal 837
2452 TACGAGTTTGCAGTGCAGTCTTCCAGCGGTGGACATGGATGGGCTTTCGGCTCTGTGGTG 2511
838 GluArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerPro 857
2512 GAGCGCTCCACCCCTGACCGTCCCTCCACACCCCTCCAGACCTGCGACTGAGCCCC 2571
858 LeuThrProSerThrValArgLeuHisTrpCysProProThrGluProAenGlyGluIle 877
2572 CTGACACCGTCCACGGTTGGCTGCTCACTGGTGGCTCCCGCCACAGAGCCCAACGGGAGATC 2631
878 ValGluTyrIleLeuLeuTyrSerAenAenHisThrGlnProGluHisGlnTrpThrLeu 897
2632 GTGGAGTATCTGATCTCTGACAGCAGCAACACACAGCAGGCTGAGCACACAGTGGACCTTG 2691
898 LeuThrThr-----GluGlyAenIlePheSerAlaGluValHisGlyLeuGluSerAsp 915
2692 CTCACCAAGCGAGGTGAGGGAAACATCTTCACTGCTCAGGTGCCATGGCTGGAGAGGCAC 2751
916 ThrArgTyrPhePheLysMetGlyValArgThrGluValGlyProGlyProPheSerArg 935
2752 ACTCGTACTTCTCAAGATGGGGCGGCACAGAGTGGAGCTGGGCTTCTCCCGC 2811
936 LeuGlnAspValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaVal 955
2812 CTGCAGGATGTGATCAGCGCTCCAGGAGAGTGTGACACTCGCTGGACATGCACTCAGTC 2871
956 ThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAla 975
2872 ACGGGCATCATCGTGGGTGTCTGCTGGGCTCTCTGCTCCCTGGGCTGTGATGTGCT 2931
976 GlyLeuArgGlnSerHisArgGluAlaLeuProGlyLeuSerSerGlyThrPro 995
2932 GGCCTGGCGGACCCCGCCACAGGGGATCCCTCCAGGCTGTCTCCACCGCCACCCCC 2991
996 GlyAenProAlaLeuTyrThrArgAlaArgLeuGlyProProSerValProAlaHis 1015

2992 GGGATATCCCGCGCTGTACTCTCAGAGCTCGGCTTGGCGCCCGCCAGCCCGCCAGTGGCCAT 3051
1016 GluLeuGluSerLeuValHisProArgProGlnAspTrpSerProProSerAspVal 1035
3052 GAATTGGAGTCCCTTGTGCACCCCGCATCCAGAGACTGGTCCCGCCAGCCCTCAGAGCTG 3111
1036 GluAspLysAlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHis 1055
3112 GAGGACAGGGCTGAACTGACAGCTTATGGGTGGCGGTGTTTCTGAAGCCCGAGTGCAC 3171
1056 SerLysArgLysIleSerTrpAlaGlnAlaGlyGlyProAenTrpAlaGlySerTrpAla 1075
3172 TCCAAAGAAAGTAAGT---GCTCAACAGCGGGCTGAGCTGGGCTGTTCTTGAGCA 3228
1076 GlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuPro 1095
3229 GGCTGTGAGTGGCCCG---GCAGGCCCGCGGCTCTGACCCCGGGCCCTGCTGCC 3285
1096 ProAlaGlyThrGlyGlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleLysSer 1115
3286 CTTGCTGGAACTGGGCAGAGCTGTTGCTGCAAGTTCTCTGCTCTGAT-----CAGGGC 3339
1116 AenGlyArgLysLysProSerProAlaCysAenGlnValGluAlaGluValIleVal 1135
3340 AATGGGAGGAGGAAGTCAACCCCGCTGAGGACAGGAGTGGAGTGAATGATGTC 3399
1136 HisSerAspPheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluPro 1155
3400 CACTCTGACTTATAGTGCATTAACGGGAAACCTGACCTCCATCTCCAGAGCTGGAGCT 3459
1156 GluGluProLeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGln 1175
3460 GAGGACCCCTGCTCCAGAGGCTCTGATCTCATCTCGGGTGTGGAGTCCAGGGCAG 3519
1176 GlyAlaAspTrpLeuGlyArgGluLeuGlyCysGlnProThrThrSerGlyProGlu 1195
3520 GGGGCGACCTGGCTGGACAGGGAGTTGGAGGGTGTGAGCTGGCAGCCCGGGCCAGAC 3579
1196 ArgLeuThrCysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnPro 1215
3580 AGATTACTGCTTGCAGAGGAGGAGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3639
1216 SerThrAlaIleGluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeu 1235
3640 GGGCAGGTGTAGAGGAGACCCCTGGAGATAGTGGCAGCTCAATCCCGCTGCGCTCTA 3699
1236 ThrValSerProSerLeuProArgAlaProValSerSerAla 1249
3700 GGAGCCAGCCCGCCAGGCTGGCCAGATCCCGGCTCTCTCTCTCT 3741

RESULT 5

US-09-754-997A-3
; Sequence 3, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2796
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (1)...(2796)

US-09-754-997A-3

Alignment Scores:

Pred. No.:	0	Length:	2756
Score:	4903.00	Matches:	932
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	74.24%	Indels:	0
DB:	10	Gaps:	0

US-09-754-997A-2 (1-1252) x US-09-754-997A-3 (1-2796)

Qy	22	GlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCysAspGluGlyProLeu	41
Db	1	GGGGAGCTGCCATTGCCCCAGGAGACAACTGCTCAAGCTGAGCTGTGATGAGGGACCCCTG	60
Qy	42	GlnValIleLeuGlyProGluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAla	61
Db	61	CAAGTGATCCTGGGCGCTTGAGCAGGCTGTGTGTCTGCACTGCACTTTGGGGGCTACAGCT	120
Qy	62	AlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGluHisGlu	81
Db	121	GCTGGGCTCCGACCAAGGGTGACATGGAGCAAGGATGGAGACACTGTACTAGAGCATATG	180
Qy	82	AsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnAsp	101
Db	181	AACCTGCACCTGCTACCCCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAAGAC	240
Qy	102	SerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTyrSerCysLeu	121
Db	241	AGCGATGATGAGAAAGCTCTTAGGATCTGGGAAGGTCACTGAGGGCAGCTATTCTCTGTCTG	300
Qy	122	AlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThrLeu	141
Db	301	GCCCAAGCCGCTAGGAGTGTGGTGGCCAGCAGAGTGTCTGTGTCAAGCTTGCACACTC	360
Qy	142	GluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArgPhe	161
Db	361	GAAGACTTCTCTCGACCCCGAGTCCCAGATTGTGGAGGAGAACGGGACAGCACGCTTT	420
Qy	162	GluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnValThr	181
Db	421	GAATGCCACACCAAGGGCTTCCAGCCGCCATCATTACTTGGGAAAAAGGACCAAGGTGACC	480
Qy	182	ValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspVal	201
Db	481	GTCCCTTGAGAGCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATGCTTAGATGTC	540
Qy	202	GlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPhe	221
Db	541	CAGACAGTGATGCAGGCTCTTACCGCTGGTGGCCACCAATTTCAGCCCGCCAAACGATTC	600
Qy	222	SerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGln	241
Db	601	AGCCAGGAGGCGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACCAGGGGCGAG	660
Qy	242	AspValValIleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnValVal	261
Db	661	GATGTGGTCAATTGTGGCAGCCCGAGAACCAACCGGTAGTGTCTCGACAGAAATGTAGTG	720
Qy	262	MetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAspGly	281
Db	721	ATGGAGTGGGTGGCCCTCTGCTGACCCACCCTTTTGTCTCTGGGTCCGACAGGATGGA	780
Qy	282	LysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAla	301
Db	781	AAGCCATCTCCACGGATGTCACTGTCTGGGCGGACCAATTACTTACTCATCGCAGCGCG	840
Qy	302	GlnProArgHisSerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPhe	321
Db	841	CAGCTTCGGCACTCTGGAGTCTATGTCTGCCAGCAACCAAGCCCTCAGCGTGACTTC	900

```
Db 1981 CGTGGAGATCAAGCTTGGAGCGTGGGGCCGTCGGCTGAAGAAAGTGAAGCAGTAT 2040
Qy
Db 702 GluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAnLys 721
Db 2041 GAATGACCCAGTAGTCCCTGGCAGGCCGTACAGGTGAGCTCTAGCTTTCACAAA 2100
Qy 722 HisGluAspGlyTyrAlaAlaValTyrLysGlyThrGluLysAlaProThrProAsp 741
Db 2101 CACGAGACCGCTACGCTGCTGTGTGAGAGGCAAGACGAGAGAGCGGCCACCGCAGAC 2160
Qy 742 LeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAnSer 761
Db 2161 CTGCTATCCAGAGGGGGCCACCGCTGCTCCCTGCCCCATGTCCACGAGAGTCAACAGC 2220
Qy 762 SerThrSerIleTyrLeuArgTyrLysLysProAspPheThrThrValLysIleValAsn 781
Db 2221 TCCACTTCCATTGGCTTCGGTGGAGAGCCAGACTTTACCACTGTCAAGATTGTCAC 2280
Qy 782 TyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThrTyrThr 801
Db 2281 TACACTGTACGCTTCGGCCCTGGGGCTCAGGAATGCTTCCCTGGTCACTACTATACC 2340
Qy 802 SerSerGlyGluAspIleLeuIleGlyGlyLeuLysProPheThrLysTyrGluPheAla 821
Db 2341 AGCTCTGGAGAGACATTTCTATTGGCGGCTGAAACCAATTTACCAAGTACGAGTTTGG 2400
Qy 822 ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr 841
Db 2401 GTAAAGTCCACGGAGTGGATATGGATGGGCCCTTTGGCTCCCTGTAGAACCTCCACC 2460
Qy 842 LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer 861
Db 2461 CTGCTGACCGGCTTCAACACCTCTCTGACCTGGCCCTGAGCCCTGACCACTCC 2520
Qy 862 ThrValArgLeuHisTyrCysProProThrGluProAsnGlyGluIleValGluTyrLeu 881
Db 2521 ACGGTTCCGTTACACTGGTGTCCCTCCCGCAGAGCCCAATGGTGGAGATTGTGGAGTATCTA 2580
Qy 882 IleLeuTyrSerAnAnHisThrGlnProGluHisGlnTyrThrLeuLeuThrThrGlu 901
Db 2581 ATTCTCTACAGCAACCAACCAACCCAGCCGGAACACAGGTGGACACTGCTCCACAGAG 2640
Qy 902 GlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLys 921
Db 2641 GGAAACATCTTCAGTCAGAGGTCATGCGCTAGAGAGTACACTCGGTATTTCTTCAAG 2700
Qy 922 MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
Db 2701 ATGGAGCGCCGACAGAGGTGGGGCTGGGGCCCTTTTCCGCTTGCAGGATGTGATTACT 2760
Qy 942 LeuGlnGluThrPheSerAspSerLeuAspValHis 953
Db 2761 CTGCAAGAGACATTTCTCAGACTCCTTGGATGTGCAC 2796
```

RESULT 6

```
US-10-239-663-28
; Sequence 28, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
```

```

; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-239-663-28

Alignment Scores:
Pred. No.: 2,348-246 Length: 2223
Score: 3167.50 Matches: 620
Percent Similarity: 90.33% Conservative: 34
Best Local Similarity: 85.64% Mismatches: 54
Query Match: 47.96% Indels: 18
DB: 12 Gaps: 2

US-09-754-997A-2 (1-1252) x US-10-239-663-28 (1-2223)

Qy 17 LeuLeuSerAlaArgGlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCys 36
Db 61 CTCCTCTCTATATGAGGGGAGCTGCTTTGGCCCCCAGGAGACGACTGTGGAGCTGAGCTGT 120
Qy 37 AspGluGlyProLeuGlnValIleLeuGlyProGluGlnAlaValValLeuAspCysThr 56
Db 121 GGAGTGGGGCCACTGCAGTGATCTTGGCCCCAGAGCGCTGCGAGTGTAACTGTATGAC 180
Qy 57 LeuGlyAlaThrAlaAlaGlyProThrArgValThrTrpSerLysAspGlyAspThr 76
Db 181 CTGGGGGCTGCTGCGCTGGACCCCGCCACAGGGTGACCTGGAGCAAGATGGGGACAC 240
Qy 77 ValLeuGluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerPro 96
Db 241 CTGCTGGAGCACACACACTTACACCTGCTGCCCAATGGTTCCCTGTGGCTGTCCAGCCA 300
Qy 97 LeuGluGlnGluAspSerAspAspGluGluAlaLeuArgIleTyrLysValThrGluGly 116
Db 301 CTAGCACCCCATGGCAGTGACAGGT--CAGTCCCTGAGGCTGTGGG--GTCAATTGANGC 357
Qy 117 SerTyrSerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValVal 136
Db 358 AACTATTCTGTGCTAGCCACCGCCCTGAGTGTCTGCCAGCCAGACTGCTGTGCTC 417
Qy 137 LysLeuAlaThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsn 156
Db 418 AAGCTTGGCCACACTCGCAGACTTCTCTCTGCACCCCGAGTCTCAGACGGTGGAGGAGAAC 477
Qy 157 GlyThrAlaArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGlu 176
Db 478 GGGACAGCTCGCTTTTGGTGGCACATTGAAGGCTGCCAGCTCCAGCTCCCATATTACTTGGAG 537
Qy 177 LysAspGlnValThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeu 196
Db 538 AAGGACCAAGTGCATGCTGAGGAGCTCGGCTCATCTGTGCTTCCCAAGCGGCTCT 597
Qy 197 GlnIleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAnSer 216
Db 598 CAGATCTCGATGTTTTCAGGAGAGTATGACGGCCCCCTACCGCTCGCTGGCCACCACTCA 657
Qy 217 AlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGlu 236
Db 658 GCTCGCCAGCACTTACGCCAGAGGCGCTTACTCAGTGTGGGCCCCACAGAGGGTCCCTGGCG 717
Qy 237 AlaThrArgGlyGlnAspValIleValAlaProGluAsnThrThrValValSer 256
Db 718 TCCACCGAGGGGAGAGACGCTGTCTATTGTGGAGCCCCCAGAGAACACACACAGTGTGTCT 777
Qy 257 GlyGlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrp 276
```



```
Db 2257 ATTAACATACACCATCGCTGTAATCTCTGTGGCTGCAGAAATGCTTCTTGGTTCTGTAC 2316
Qy 800 TyrThrSerSerGlyGluAuplleLeulleGlyLeuLysProPheThrLysTyrrGlu 819
Db 2317 CTTCAAAATCAGAAATCTCATGTTGGTTCAAGGCTTAGAACCAACACCAATACGAA 2376
Qy 820 PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg 839
Db 2377 TTTGCCGTCGATACATGATGATCAGCTTTCAGTCTTCCAGTCTTGGAGCCCTGTAGTCTACCAT 2436
Qy 840 SerThrLeuProAspArgProSerThrProSerAspLeuArgLeuSerProLeuThr 859
Db 2437 TCTACTCTTCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2496
Qy 860 ProSerThrValArgLeuHisThrCysProProThrGluProAsnGlyGluValGlu 879
Db 2497 GATGACACTGCGCTTCTTGGAAACCCCTGATGCCAGAAAACAGTTGTGCCCGC 2556
Qy 880 TyrLeulleLeuTySerAsnAsnHisThrGlnProGluHisGlnTrrThrLeuLeuThr 899
Db 2557 TATACTATCTATATGATCTAGGAGGCTGGATTGCGAGAGATGCGAGGCTTACAC 2616
Qy 900 ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrrPhe 919
Db 2617 CGTGAAGGGGCAATAACCATGCTTGTCTAGAAAACCTTGTAGCAGGAAATGTGTACATT 2676
Qy 920 PheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
Db 2677 GTCAAGATATCTGCATCAATGAGGTGGGAGAGGACCCCTTTTCAAAATCTGTGGAGCTG 2736
Qy 940 IleThrLeuGlnGluThrPheSerAspSer----- 949
Db 2737 GCAGTACTTCCAAAGGAACTCTGTAATCAATCAGAGGCCAAGCGTTAGATTCTGCT 2796
Qy 950 -----LeuAspValHisAlaValThrGlyIle 959
Db 2797 GATGCCAAAGTTATTTCAGGATATTACCATCTGCAGCAAAAATCAATGATCTGGCATTTGCT 2856
Qy 960 ValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln 979
Db 2857 GTAGGTGTGGCAGGCTTGACCTGCATCTCATCTGTGTCTCATCTTGTGATATACCGA 2916
Qy 980 SerSerHisArgGluAla-----LeuPro 987
Db 2917 AGTAAGCCAGGAATCATCTGCTTCCAGAGCGGCAGAGATGGAATCAACAGTTACCT 2976
Qy 988 GlyLeuSerSerSerGlyThrProGlyAsn 997
Db 2977 CGTACCAGTGCCTCTTACGTAGTGGAAAT 3006
```

RESULT 8

```
US-10-231-353-1
; Sequence 1, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20030023064A1el Human Phosphatases and Polynucleotides Encod
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231.353
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3210
; TYPE: DNA
```

```
ORGANISM: homo sapiens
US-10-231-353-1
```

Alignment Scores:

```
Pred. No.: 2,25e-131 Length: 3210
Score: 1746.50 Matches: 397
Percent Similarity: 54.95% Conservative: 169
Best Local Similarity: 38.54% Mismatches: 381
Query Match: 26.45% Indels: 83
DB: 14 Gaps: 18
```

```
US-09-754-997A-2 (1-1252) x US-10-231-353-1 (1-3210)
```

```
Qy 9 GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln 28
Db 43 GGGATGCTGCTCGCGCGCTCTCTGCTCTCTGCTGCTCAGT-----CCTTTGCCAGGA 96
Qy 29 GluThrThrVal---LysLeuSerCysAspGlyProLeuGlnValLeuLeuGlyPro 47
Db 97 GTGTGTGTCTTTAGCGAATCTGCTTTTGTAAAGAACACACAGGATGTAACTGTACACAAGA 156
Qy 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArg 67
Db 157 AAGGACCCAGTCTGTTTAGATTCCAGGCTCAGCGAGAGATT-----CCTATTAAAG 207
Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuLeuHisGluAsnLeuHisLeuLeuPro 87
Db 208 GTCAATGTTGAAATATGGAGCAAAATGTCTGMAAATAAAGCGATCGAGGTTCTTCT 267
Qy 88 AsnGlySerLeuTrpLeuSerSerProLeuGlnGlnGluAspSerAspGluGluAla 107
Db 268 AACGGCTCTTTATACATCAGT-----GAGGTGGAAGCGAGGCGAGGAGCAGTCC 318
Qy 108 LeuArgIleTrpLysValThrGluGlySerTyrrSerCysLeuAlaHisSerProLeuGly 127
Db 319 -----GATGAAGGATTATTACGTGCTTGGCAATGAACAATAATATGGA 360
Qy 128 ValValAlaSerGlnValAlaValValLeuLeuAlaThrLeuGluAspPheSerLeuHis 147
Db 361 GCATTTCTTAGTCAAAAGCTCATCTGCTTATCACTATTTCTGCAATTTGAAGTCCAG 420
Qy 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
Db 421 CCAATTTCCATCAGGTCCAGGAAGGTGGAGTTGCTCGATTTCGATGCAAGATTTTCATCC 480
Qy 168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
Db 481 CACCCTCTCGACGTATACATGGGAGTTTCAATCGGACAACTTACCTATGACTATGGAC 540
Qy 187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
Db 541 AGGATTAATCTGCTTCCACAGGAGTATTGCGATCTATGATGTTCAGCCAAAGGGATTCT 600
Qy 207 GlySerTyrrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
Db 601 GGAATATTCTGTTGTTGCTGCGACCTGTAGCCACCGACGCTAAAGTATGAGAGGCTCG 660
Qy 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIleVal 246
Db 661 CTAATCTGTGATT-----CCAGCTTAAGGAGTCAAAATCTTCCACACACCAACCAATTTATA 714
Qy 247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
Db 715 GCAGGTCCACAGAACATAACACATCTCTTCATCAGACTGTAGTTTGGATATGATGCGCC 774
Qy 267 SerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThr 286
Db 775 ACAGGAATCCCAAAACCAATCATTTCTTGGAGCGGCTTGTATCACAATATCCATTGATGTC 834
Qy 287 ---AspValIleValLeuGlyArgThrAsnLeuLeuAlaSerAlaGlnProArgHis 305
Db 835 TTTAATACTCGGCTACTTTGGAAATGGTAATCTCATGATATCTCATGATATCTGATGTGAGCTACAACAT 894
```


Qy	187	ArgLeuIleThrLeuProIysTrpTrpLeuLeuGlnIleLeuAspValGlnAaspSerAspAla	206
Db	731	AGGATAACTGCCCTACCAACAGGAGTATTGCAGATCTATGATGTCAGCCAAAGGGATTCT	790
Qy	207	GlySerTrpArgCysValAlaThrAenSerAlaArgGlnAArgPheSerGlnGluAlaSer	226
Db	791	GGAAATATTGTTGTTATGCTGCCACTGTAGGCCACCGAGCTAAATAGTATGGAGGGCTCG	850
Qy	227	LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAaspValValIleVal	246
Db	851	CTAACTGTGATT-----CCAGCTAAGGAGTCAAAATCCTTCACACACACARCAATTATA	904
Qy	247	AlaAlaProGluAenThrTrpValValSerGlyGlnAasnValValMetGluCysValAla	266
Db	905	GCAGGTCCACAGAAATAACAACATCTCTTCATCAGACTGTAGTTTGGAAATGCATGGCC	964
Qy	267	SerAlaAaspProThrProPheValSerTrpValArgGlnAaspGlyIysProIleSerThr	286
Db	965	ACGAAAAATCCCAACCAATCTTCTTCGGAGCGGCTTGAATCACAANTCCATTGATGTC	1024
Qy	287	---AaspValIleValLeuGlyArgThrAenLeuLeuIleAlaSerAlaGlnProArgHis	305
Db	1025	TTTAATACTCGGTACTTGGAAATGTTATCTCATGATATCTGATGTCAGGCTACAACT	1084
Qy	306	SerGlyValTrpValCysArgAlaAenLysProLeuThrArgAaspPheAlaThrAlaAla	325
Db	1085	GCTGGAGTATATGTTTGGGCCACTACCCCTGGCACACGCCAACTTACAGTTGCTATG	1144
Qy	326	AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg	345
Db	1145	GCACCTTTAACTGATTTAGCTCCTCTTCATTTGTTGTAATGGCCAGAAAGTTTAAACA	1204
Qy	346	ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu	365
Db	1205	CCTCGAGCTGGCACTGCTCGATTGTGTGTCAGGCGAGAGGANTCCCTCTCCAGAGTG	1264
Qy	366	HisTrpLeuHisAaspGlyIleProLeuArgProAasnGlyArgValIysValGlnGly	385
Db	1265	TCATGGTTGAAAAATCGAGGAAGATACATTCGAATGGTAGAATTAATGATGAC-----	1318
Qy	386	GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAaspAlaGlyTrpTrpGlnCysVal	405
Db	1319	AACAGTAAATTTGTTAAATTAACCAAGTATTTCCTCGAAGATGATGCTATTTATCATG	1378
Qy	406	AlaGluAenSerAlaGlyThrAlaCysAlaAlaAlaProLeuAlaValValValArgGlu	425
Db	1379	GCTGAGAAATAGCCAGAGATCTATTTATCTAGAGCCAGACTGACTGTAGTGTGTGAA	1438
Qy	426	GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu	445
Db	1439	GACAGACCCAGTCTCCCTATAATGATACATGCTGAAACCATGTCAAGCTCAGCCATTCTT	1498
Qy	446	ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTrp	465
Db	1499	TTAGCCTGGGAGAGGCCACTTTTATAATTCAGACAAAGTCAATGCTATTCTGTACACTAC	1558
Qy	466	GlnIleAlaArgGlyValAaspAenValGluTrpGlnPheAlaValAaenAenAaspThr	485
Db	1559	ATGAAGACAGAGGTTTAAATATGAAGAGTATCAAGTAGTCTCGGAAATGACACAAC	1618
Qy	486	GluLeuGlnValArgAaspLeuGluProAasnThrAspTrpGluPheTrpValValAlaTrp	505
Db	1619	CATTATATTATGATGACTAGAGCCTGCCAGCAATATATCTTCTACATTTGTAGCATAT	1678
Qy	506	SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAaspVal	525
Db	1679	ATGCCAATGGGACGACCATGTTCTGACCATGTGACACAGTAATCTCTAGAGGATGTT	1738
Qy	526	ProSerAlaAlaProGlnLeuThrLeuSerSerProAasnProSerAaspIleArgValAla	545
Db	1739	CCCTTGAGACCTCTGAAATAGTTTGACAAGTCGAAGTCCCACTGATATTTCACTCTCC	1798

Db 100 CAGTGTGGCAATGAACAATAATATGAGCCATTCTTAGTCAAAAAGCTCATCTTGCCCTTA 159
Qy 139 AlathrLeuGluAaspPheSerLeuHisProGluSerGlnIleValGluGluAasnGlyThr 158
Db 160 TCAACTATTCTTGCAATTTGAAGTCCAGCCAAATTTCCACTGAGGTCCAGGAAGTGGAGTT 219
Qy 159 AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleThrTrpGluLysAasp 178
Db 220 GCTCGATTGTCATCAAGATTTCATCCACCCTCTCCAGTCATAAATGGAGTTCAAT 279
Qy 179 GlnValThrValPro---GluGluProArgLeuIleThrLeuProLysTrpLeuGln 197
Db 280 CGGCAAACTCTACCTATGACTATGAGCAGGATAACTGCCCTTACCACACAGGAGTATTGCG 339
Qy 198 IleLeuAaspValGlnAaspSerAaspAlaGlySerTrpArgCysValAlaThrAasnSerAla 217
Db 340 ATCTATGATGTCAGCCAAAGGATTCGGAATATTCGTTGATTTGCTGCCACTGTAGCC 399
Qy 218 ArgGlnArgPheSerGlnGluAaspSerLeuThrValAlaLeuAaspGlySerLeuGluAla 237
Db 400 CACCGAGTAAAGTATGAGGCTCGCTAACTGTGATT-----CCAGCTAAGGAGTCA 453
Qy 238 ThrArgGlyGlnAaspValIleValAlaAlaProGluAasnThrThrValValSerGly 257
Db 454 AAATCCTTCCACACACCAACCAATATATAGCAGGTCCACAGAACATAACAACATCTCTTCAT 513
Qy 258 GlnAasnValValMetGluCysValAlaSerAlaAaspProThrPheValSerTrpVal 277
Db 514 CAGACTGTAGTGTGGTAATGATGCGCCACAGGAATCCCAACCAATCTTCTTTGGAGC 573
Qy 278 ArgGlnAaspGlyLysProIleSerThr---AaspValIleValLeuGlyArgThrAasnLeu 296
Db 574 GCGCTTGATCAAAATCCATTTGATGTCTTAAATCTCGGTACTTGGAAATGGTAATCTC 633
Qy 297 LeuIleAaspAlaGlnProArgHisSerGlyValTyrValCysArgAlaAasnLysPro 316
Db 634 ATGATATCTGATGTGAGGTACCAACATGCTGGAGTATATGTTGTCGGGCCACTACCCCT 693
Qy 317 LeuThrArgAaspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle 336
Db 694 GGCACGCAACTTTTACATGTTGCTATGGCACTTTAACTGATTAGTCTCTCTTCATTT 753
Qy 337 SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaAaspPheValCysArg 356
Db 754 GTTGAATGCCAGAAAGTTAAACAGGCTCGAGCTGCATCTCGATTTGTGTGTGTCAG 813
Qy 357 AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAaspGlyIleProLeuArgPro 376
Db 814 GCAGAAAGGAATCCCTCTCCCAAGATGTCTGTTGAAATAATGGAAGAAGATACATTCG 873
Qy 377 AasnGlyArgValIleValGlnGlyGlySerLeuValIleThrGlnIleGlyLeu 396
Db 874 AATGGTGAATTAATAATGTAC-----AACAGTAAATTTGGTAATTAACCAATTTATTCCT 927
Qy 397 GlnAaspAlaGlyTyrTrpGlnCysValAlaGluAasnSerAlaGlyThrAlaCysAlaAla 416
Db 928 GAAGATGATGCTATTATTCAGTGAATGCTGAGATAGCCAGGATCTATTATTATCTAGA 987
Qy 417 AlaProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThrAla 436
Db 988 GCCAGACTGACTGTAGTATGTCAGAGACAGACAGCCAGTCTCCCTATAATGTATCATGCT 1047
Qy 437 ThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlu 456
Db 1048 GAAACCACTGTCAGCTCAGCAATTTCTTAGCTGGGAGAGGCCATTTTATATTTCAGAG 1107
Qy 457 GlnIleIleGlyPheSerLeuHisTrpGlnLysAlaArgGlyValAaspAasnValGluTyr 476
Db 1108 AAAGTCATGCTATTCTGTACACTACATGAAGCAGAGAGTTTAAATTAATGAAGATAT 1167
Qy 477 GlnPheAlaAasnAasnAaspThrThrGluLeuGlnValArgAaspLeuGluProAasnThr 496
Db 1168 CAAGTAGTCATCGGAAATGACACAACTCATTTATATTGATGACTTAGAGCCTGCCAGC 1227

Qy 497 AaspTyrGluPheTyrValValAlaTatYrSerGlnLeuGlyAlaSerArgThrSerSerPro 516
Db 1228 AATTATATCTTCTACATTTAGCATATATGCCAATCGGAGCCAGCCAGATGCTGACCAT 1287
Qy 517 AlaLeuValHisThrLeuAaspValProSerAlaAlaProGlnLeuThrLeuSerSer 536
Db 1288 GTGACACAGAAATACTCTAGAGGATGTTCCCTCGAGACCTCTCGAAAATTAGTTTGACAAGT 1347
Qy 537 ProAasnSerAaspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAasnGly 556
Db 1348 CGAAGTCCCACTGATATTCTATCTCTGCTGCCAATCCAGCCAAATATATCGCGGGGC 1407
Qy 557 GlnValLeuLysTyrLysIleGluTyrGlyLeuGlyLysGluAaspGlnValPheSerThr 576
Db 1408 CAAGTGTGCTGTATCGTTGCTTTCGCTTAAGTACTGAGAAATTCATTCCAAGTTCTG 1467
Qy 577 GluValProGlyAasnGluThrGlnLeuThrLeuAasnSerLeuGlnProAasnLysValTyr 596
Db 1468 GAGCTCCGGGACCAACGATGAGTACTCTTTTGAAGGCTCGAAACCTGACAGTGTCTAC 1527
Qy 597 ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln 616
Db 1528 CTGGTTGGAATTAATCTCTGCCACAGAGTGGGCTCGGAGAGTCACTAGTATGACTTCA 1587
Qy 617 HisArgThrProGlyValHisAasnGlnSerHisValProPheAlaProAlaGluLeuLys 636
Db 1588 CATAGAGCGCCCAAAAGCT---ACAAGCGTGAAGCCCTTAAGTCTCA---GAGTTGCAAT 1641
Qy 637 ValArg---AlaLysMetGluSerLeuValSerTrpGlnProProHisProThr 655
Db 1642 TTGGAGCCTCTGAACCTGTACCACCACTTCTGTGAGGTGGCAGCAAGATGTAGAGGACACA 1701
Qy 656 ---GlnIleSerGlyTyrLysLeuTyrTrpGlyValGlyThrGluGluAlaAasp 674
Db 1702 GCTGCTATTGAGGCTACAGCTGTACTACAGAAAGAGGCGCAGCAGAGAAAT----- 1755
Qy 675 GlyAaspArgProGlyGlyArgGlyAaspGlnAlaTrpAaspValGlyProValArgLeu 694
Db 1756 -----GGGCCCATTTTCTTG 1770
Qy 695 LysLysLysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluVal 714
Db 1771 GATACCAAGGACCTACTCTATCTCTCAGTGGCTTAGACCCCAAGAAATAATCATCATGTG 1830
Qy 715 LysLeuValAlaPheAasnLysHisGluAaspGlyTyrAlaAlaValTrpLysGlyLysThr 734
Db 1831 AGACTCTGGCTTACAAACAACATAGACGATGGCTATCAGGCA-----GAT 1875
Qy 735 GluLysAlaProThrProAaspLeuProIleGlnArgGly-----ProProLeu 750
Db 1876 CAGACTGTACGACTCCAGGATCGGTGTCTGTTGCTGATCGATGGTCCCTCTCCACCA 1935
Qy 751 ProProAlaHisValHisAlaGluSerAasnSerThrSerIleTrpLeuAsgTrpLys 770
Db 1936 CCACCCCAACCTCTCTATGCGAAGGCTAACACCTCATCTTCCATCTTCTCTGCACTGAGG 1995
Qy 771 LysProAaspPheThrValLysIleValAasnTyrThrValArgPheGlyProTyrGly 790
Db 1996 AGGCTGCAATTCACCGCTGCACAAATCAATTAATACACCTCCGCTGTAAATCTGTTGGC 2055
Qy 791 LeuArgAasnAlaSerLeuValThrTyrTyrThrSerSerGlyGluAaspIleLeuIleGly 810
Db 2056 CTGAGAGATGCTTTTGGTTCTGTACTTCAAAACATCAGAAATCATATGTTGGTTCAA 2115
Qy 811 GlyLeuLysProPheThrLysTyrGluPheAlaValGlnSerHisGlyValAaspMetAasp 830
Db 2116 GGTCTAGAACCAACACCAAAATACGAATTTGCGGTTTCGATTACATGTGATGATCAGCTTTC 2175
Qy 831 GlyProPheGlySerValValGluArgSerThrLeuProAaspArgProSerThrProPro 850
Db 2176 AGTCTTTGGAGCCCTGTAGTCTACCAATTTACTCTTCTTCAGAAAGCACCGAGCGGCCCA 2235

1048	DB	GAACCATGTC	AAGCTCAGCC	ATCTTTTAGCCTCGGAGAGGCC	ACATTTTATAATTCAGAC	1107
457	QY	GlnIleIleGlyPheSerLeu	HisTyrGln	LeuAlaArgGlyVal	AspAsnValGluTyr	476
1108	DB	AAAGTCATTC	CGCTATTCTGTAC	ACTACATCA	TGAAAGCAGAGGTTTAAATTAATGAAGAGTAT	1167
477	QY	GlnPheAlaValAsnAsn	AspThrThr	GluLeuGlnVal	ArgAspLeuGluProAsnThr	496
1168	DB	CAAGTAGTCAT	CGGAATGACAC	AACTCATATTATTAT	TGATGATCACTTAGCAGCCTCGCAGC	1227
497	QY	AspTyrGluPheTyrVal	ValAlaTyrSerGln	LeuGlyAlaSerArgThr	SerSerPro	516
1228	DB	AAATTA	TACTTTCATCTTGAT	GCATATATCCCAATGGAGCCAGC	AGATGCTTGACCAT	1287
517	QY	AlaLeuValHisThr	LeuAspValProSer	AlaAlaProGln	LeuThrLeuSerSer	536
1288	DB	GTGACACAGA	ATACTCTAGAGGATG	TTCCTCGAGACCTCTCTGAAATTAGTTTGACAAGT	1347	
537	QY	ProAsnProSer	AspIleArgValAla	TrpLeuProLeuProSer	SerLeuSerAsnGly	556
1348	DB	CGAAGTCCCA	CTGATATTTCTCATCT	CTCGTGCATCCCAATCCAGCCAAATAT	ATCCGCGCGGGC	1407
557	QY	GlnValLeuIle	ysTyrIleGluTyrGly	LeuGlyIysGluAspGln	ValPheSerThr	576
1408	DB	CAAGTGGTGT	GTATCGCTTGTCTTC	CCGCTTAAGTACTGAGAATCAAT	TCCCAAGTTCTG	1467
577	QY	GluValProGly	AsnGluThrGln	LeuThrLeuAsnSerLeu	GlnProAsnIysValTyr	596
1468	DB	GAGCTCCCGGG	ACCACGCATGAGTACCTTT	TGGAGGCTTGAACCTGACAGTGCTAC	1527	
597	QY	ArgValArgIle	SerAlaGlyThrGly	AlaGlyTyrGlyValProSer	GlnTrpMetGln	616
1528	DB	CTGCTTCGGAT	TACTGTGTGCAC	CAGAGTGGGGCTGGGAGAGTCATCAGTAT	TGCAGCTTCA	1587
617	QY	HisArgThrProGly	ValHisAsnGlnSerHis	ValProPheAlaProAla	GluLeuIys	636
1588	DB	CATAGGACG	CCCCAAAGCT--ACA	AGCGTGAAGCCCTTAAGTTC	CA--GAGTTGCAT	1641
637	QY	ValArg--Ala	IysMetGluSerLeu	ValValSerTrpGlnProPro	HisProThr	655
1642	DB	TTGGAGCCTCT	GAACTGTACACCA	CCATTTCTGTGAGGTGGCAGCAGATGTAGAGGACACA	1701	
656	QY	---Gln	IleSerGlyTyrIysLeu	TyrTrpGlyValGlyThrGlu	GluAlaAsp	674
1702	DB	GCTGCTATT	CAGGGCTACAAGCTGTACTACA	GAGAGAGGCGCAGCAGAGAAT----	1755	
675	QY	GlyAspArgPro	ProGlyGlyArgGlyAsp	GlnAlaTrpAspValGlyPro	ValArgLeu	694
1756	DB	-----	-----	-----GGGCCCAATTTTCTTG	1770	
695	QY	LysIysIysVal	LysGlnTyrGluLeuThrGln	LeuValProGlyArgProTyrGlu	Val	714
1771	DB	GATACCAAG	AGACCTACTCTATCTCTCAGT	TGGCTTAGACCCAGAGAAATATCATGTG	1830	
715	QY	LysLeuValAla	PheAsnIysHis	GluAspGlyTyrAlaAlaValTrp	LysGlyLysThr	734
1831	DB	AGACTCTCG	CTGTACAACAACATAGACGATGGCTAT	TCAGCA-----GAT	1875	
735	QY	GluIysAlaPro	ThrProAspLeuProIle	GlnArgGly-----Pro	ProLeu	750
1876	DB	CAGACTGT	CAGCATCCAGGATGCGTCTGT	TGTGATCGCATGGTCCCTCTCTCCACCA	1935	
751	QY	ProProAlaHis	ValHisAlaGluSerAsnSer	ThrThrSerIleTrpLeu	ArgTrpLys	770
1936	DB	CCACCCCA	CCATCTCTATGAGAGGCTAA	CACCTCATCTTCCATCTTCTTCGCACTG	GAGG	1995
771	QY	LysProAspPhe	ThrThrValLysIle	ValAsnTyrThrValArgPhe	GlyProTrpGly	790
1996	DB	AGGCTGCAT	TCCCGCTGCACAAATCATTA	ACTACACCATCCGCTGTAACTCCTGT	TGGC	2055
791	QY	LeuArgAsnAla	SerLeuValThrTyrTyrThr	SerSerGlyGluAspIleLeu	IleGly	810
2056	DB	CTGCAGAA	TGCTCTTTGGTTCTGTACTCTT	CAAACTCAGAACTCACA	TGTGGTCA	2115

Qy	811	GlyLeuLysProPheThrIstyTyrGluPheAlaValGlnSerHisGlyValAspMetAsp	830
Db	2116	GGTCTTGAACCAACACCAATACGAATTTGCGTTGCAATTATCATGTGGATCAGCTTTCC	2175
Qy	831	GlyProPheGlySerValValGluArgSerThrLeuProAspArgProSerThrProPro	850
Db	2176	AGTCCTTGGAGCCCTGTAGTCTACCATCTTCTCTCCAGAACCAACGAGCCCAACA	2235
Qy	851	SerAspLeuArgLeuSerProLeuThrProSerThrValArgLeuHisTrpCysProPro	870
Db	2236	GTTGGAGTAAAGTGACATTAATAGAGGATGACACACTGCCCTGGTTCTTGGAAACCCCT	2295
Qy	871	ThrGluProAsnGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGln	890
Db	2296	GATGGCCCAAGAAAGTTGTGACCCGCTATATCTTATATGCACTTAGGAAGGCGCTGG	2355
Qy	891	ProGluHisGlnTrpThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHis	910
Db	2356	ATTGCAGGAGTGGCAGGCTTTACACCGTGAAGGGCAATAACCATGGCTTTGCTAGAA	2415
Qy	911	GlyLeuGluSerAspThrArgTyrPhePheLysMetGlyAlaArgThrGluValGlyPro	930
Db	2416	AACTTGGTACGACGAAATGTGTACATTTGTTCARATATCTGCATCCCAATGAGGTGGAGAA	2475
Qy	931	GlyProPheSerArgLeuGluAspValIleThrLeuGlnGlnThrPheSerAspSer---	949
Db	2476	GGACCCCTTTCAATCTCTGGGAGCTGGCAGTACTTCCAAAGGAAACCTCTGAAATCAAAT	2535
Qy	950	-----Leu 950	
Db	2536	CAGAGGCCCAAGCGTTTAGATTCTGCTGATGCCAAAGTTTATTTCAGGATATTACCATCTG	2595
Qy	951	AspValHisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuLeu	970
Db	2596	GACCAAAATCAATAGCTGTCATCTGCTGAGGTGTGGCATAGCTTACCTTGATCCTC	2655
Qy	971	AlaCysMetCysAlaGlyLeuArgGlnSerSerHisArgGluAla	985
Db	2656	ATCTGTGTTCTCATCTTGATATACCGAAGTAAAGCCAGGAAATCATCTGCTTCCAAGACG	2715
Qy	986	-----LeuProGlyLeuSerSerSerGlyThrProGlyAsn 997	
Db	2716	GCACAGAATGGAACCTCAACAGCTTACCTCGTACAGTGCCTCTCTAGCTAGTGGAAAT	2772
RESULT 12			
US-09-754-997A-5			
; Sequence 5, Application US/09754997A /			
; Patent No. US20020102551A1			
; GENERAL INFORMATION:			
; APPLICANT: Salbaum, Michael J.			
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic			
; TITLE OF INVENTION: Acids and Methods of Use			
; FILE REFERENCE: P-NI 4552			
; CURRENT APPLICATION NUMBER: US/09/754, 997A			
; CURRENT FILING DATE: 2001-01-04			
; PRIOR APPLICATION NUMBER: US 60/174,496			
; PRIOR FILING DATE: 2000-01-04			
; PRIOR APPLICATION NUMBER: US 60/205,769			
; PRIOR FILING DATE: 2000-05-19			
; NUMBER OF SEQ ID NOS: 45			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5			
; LENGTH: 825			
; TYPE: DNA			
; ORGANISM: Mus musculus			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1) ... (825)			
US-09-754-997A-5			

Alignment Scores:
Pred. No.: 6.19e-110 Length: 825

Db 661 CTAAGTGTGATT-----CCAGCTAAGGAGTCAAAATCTCTCCACACACACCAACCAATTATA 714
Qy 247 AlalaProGluAenThrThrValValSerGlyGlnAenValValMetGluCysValala 266
Db 715 GCAGGTCACAGACATACACACATCTCTTCATCAGACTGTAGTTTTGGAAATGATGCGCC 774
Qy 267 SerAlaAspProThrProPheValSerTyrValArgGlnAspGlyLysProIleSerThr 286
Db 775 ACAGGAATCCCAACCAACATCATTTCTTGGAGCGCCTTGATCACAATCCATTCATGATGC 834
Qy 287 ---AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
Db 835 TTTAATACTCGGGTACTTGGAAATGTAATCTCATGATATCTCATGATCGAGGCTACAAACAT 894
Qy 306 SerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
Db 895 GCTGGAGTATATGTTTGGCGGCCACTACCCCTGGCACACCAACTTTACAGTTGCTATG 954
Qy 326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
Db 955 GCACTTTAATGTTATGATCTCTCTTCATTTGTTGATGGCCAGAAAGTTTAAACAGG 1014
Qy 346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
Db 1015 CCTCGAGCTGGCACTGCTCGATTTGTGTGTGTCAGGACAGAAATCCCTCTCCCAAGATG 1074
Qy 366 HisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGlyGly 385
Db 1075 TCATGGTTTGAATAATGGAAAGATACATTCGAATGGTAGAATTTAAATGTATC----- 1128
Qy 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysVal 405
Db 1129 AACAGTAAATTTGGTAATTAACAGATTTATCTCTGAGATGATGCTATTATTCAGTGCATG 1188
Qy 406 AlaGluAenSerAlaGlyThrAlaCysAlaAlaAlaProLeuAlaValValValArgGlu 425
Db 1189 GCTGAGATAGCCAAAGGATCTATTTTACTAGAGCCAGACTGACTGTAGTGTATGATGACAA 1248
Qy 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu 445
Db 1249 GACAGACCCAGTGTCTCTATATGATGCTGAAACCATGTCTCAAGCTCAGCCATCTT 1308
Qy 446 ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr 465
Db 1309 TTAGCCTGGAGAGCCACTTTTATATTCAGACAAAGTCATTCCTATCTGTACACTAC 1368
Qy 466 GlnLysAlaArgGlyValAspAenValGluTyrGlnPheAlaValAsnAsnAspThrThr 485
Db 1369 ATGAAGCAGAGGTTTAAATAATGAAGATATCAAGTAGTATCATCGGAATGACACAACT 1428
Qy 486 GluLeuGlnValArgAspLeuGluProAenThrAspTyrGluPheTyrValValAlaTyr 505
Db 1429 CATATATATTGATGACTTAGAGCTCGCCAAATTTATCTTTCTACATTTGTAGCATAT 1488
Qy 506 SerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHisThrLeuAspAspVal 525
Db 1489 ATGCCAATGGAGGAGCCAGATGCTGACCATGTGACACAGAAATACTCTAGAGATGAC 1548
Qy 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
Db 1549 CCCAGA----- 1554
Qy 546 TrpLeuProLeuProSerSerLeuSerAenGlyGlnValLeuLysTyrLysIleGluTyr 565
Db 1554 ----- 1554
Qy 566 GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
Db 1554 ----- 1554
Qy 586 ThrLeuAenSerLeuGlnProAenLysValTyrArgValArgIleSerAlaGlyThrGly 605
Db 1554 ----- 1554

Qy 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAenGln 625
Db 1554 ----- 1554
Qy 626 SerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuVal 645
Db 1554 ----- 1554
Qy 646 ValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrpGly 665
Db 1554 ----- 1554
Qy 666 GluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyArgGlyAspGln 685
Db 1554 ----- 1554
Qy 686 AlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyrGluLeuThrGln 705
Db 1554 ----- 1554
Qy 706 LeuValProGlyArgProTyrGluValLysLeuValAlaPheAenLysHisGluAspGly 725
Db 1555 -----AGAAATATCATGTGAGACTCTCGCTTACAAACACATAGACGATGGC 1602
Qy 726 TyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuProIleGln 745
Db 1503 TATCAGGCA-----GATCAGACTGTACAGACTCCAGGATGGTCTGT 1647
Qy 746 ArgGly-----ProProLeuProAlaHisValHisAlaGluSerAenSer 761
Db 1648 CGTGATCGCATGGTCCCTCTCCACACACCCACCATCTCTATGCGAAGGCTAAACACC 1707
Qy 762 SerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsn 781
Db 1708 TCATCTCCATCTCTCCACTCGAGAGCGCTGCATTCCCGCTGCACCAATCATTAAC 1767
Qy 782 TyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyrThr 801
Db 1768 TACACCATCGCTGTAATCTCTGTGGCTGCGAGATGCTCTTTGGTTCTGTACTTCAA 1827
Qy 802 SerSerGlyGluAspIleLeuIleGlyLysLysProPheThrLysTyrGluPheAla 821
Db 1828 ACATCAGAACTCACAATGTTGGTTCAAGTCTAGAACCAACCAACCAATACGAATTTGCC 1887
Qy 822 ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr 841
Db 1888 GTTCGATTATCATGTGGATCAGCTTTCAGTCTCTTGAGCCCTGTAGTCTACCATTTCTACT 1947
Qy 842 LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer 861
Db 1948 CTTCCAGAGCACCAGCAGCGCCACCATGTTGGAGTAAAGTGACATTAATAGAGATGAC 2007
Qy 862 ThrValArgLeuHisTyrCysProProThrGluProAenGlyGluIleValGluTyrLeu 881
Db 2008 ACTGCCCTGGTTCTTGAAACCCCTGATGCGCCCAAGAACAGTTGTGACCCGCTACTACT 2067
Qy 882 IleLeuTyrSerAenAenHisThrGlnProGluHisGlnTrpThrLeuLeuThrThrGlu 901
Db 2068 ATCTTATATGATCTAGGAAGCGCTGGATTTCAGAGAGATGCGAGGCTCTTACACCGTGA 2127
Qy 902 GlyAenIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLys 921
Db 2128 GGGCAATAACCATGGCTTCTGTAGAAAACCTTGTAGCAGGAATGTGTATCTGTCAAG 2187
Qy 922 MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
Db 2188 ATATCTGCATCAATGAGTGGAGAGGACCCCTTTCAATTCGTGTGGAGCTGCACATA 2247
Qy 942 LeuGlnGluThrPheSerAspSer----- 949
Db 2248 CTTCCAAAGGAAACCTCTGTAATCAATCAGAGGCCCAAGCGTTTAGATTCTGTCTGATGCC 2307

```
QY 950 -----LeuAspValHisAlaValThrGlyIleIleValGly 961
Db 2308 AAGTTTATTACAGATATTACCATCTGGACCAAAATCAATGACTGGCATTTGCTAGGT 2367
QY 962 ValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSerSer 981
Db 2368 GTTGGCATAGCTTGGATGATGCTGCTCATCTGTTCTCATCTTGTATATACCGAGTAA 2427
QY 982 HisArgGluAla -----LeuProGlyLeu 989
Db 2428 GCCAGCAAAATCATCTGCTTCCAAAGCGCACAGATGGAATCAACAGTTACCTCGTACC 2487
QY 990 SerSerSerClyThrProGlyAsn 997
Db 2488 AGTGCCTCTTAGCTAGTGGAAAT 2511

RESULT 14
US-10-231-353-9
; Sequence 9, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20030023064A1 Human Phosphatases and Polynucleotides Encod
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231,353
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-231-353-9

Alignment Scores:
Pred. No.: 3,89e-102 Length: 2958
Score: 1384.00 Matches: 334
Percent Similarity: 46.11% Conservative: 140
Best Local Similarity: 32.49% Mismatches: 310
Query Match: 20.96% Indels: 244
DB: 14 Gaps: 14

US-09-754-997A-2 (1-1252) x US-10-231-353-9 (1-2958)
QY 9 GlyLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln 28
Db 43 GGGATGCTGCTCGCGCGCTCTGCTCTGCTGCTCAGT-----CCTTTGCCAGGA 96
QY 29 GluThrVal---LysLeuSerCysAspGluGlyProLeuGlnValIleLeuGlyPro 47
Db 97 GTGTGTGCTTTAGCGAAGCTCTTTTGTAAAGAACCCACAGGATGTAATGTCACAGA 156
QY 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProThrArg 67
Db 157 AAGGACCCAGTCGTTTGTAGTCCAGGCTCACGGAGAGT-----CCTATTAG 207
QY 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuPro 87
Db 208 GTCACATGGTTGAAAAATGGACAAAAATGCTGAAAAATAAACGGATCGAGGTTCTTCT 267
QY 88 AsnGlySerLeuTrpLeuSerProLeuGluGlnGluAspSerAspGluGluAla 107
Db 268 AACGGCTCTTTATACATCAGT-----GAGGTGGAAGGCGGAGGAGGAGTCC 318
QY 108 LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly 127
```

```
Db 319 -----GATGAAGGATTTTATCAGTGTCTGGCAATGAAACAATATGGA 360
QY 128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
Db 361 GCCATTTAGTCAAAAGCTCATCTGCTTATCACTATTCTTCATTTGAGTCCAG 420
QY 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaAspPheGluCysHisThrLysGly 167
Db 421 CCAATTTCCACTGAGTCCACGAAGGTGAGTTGCTCGATTTGTCATGCAAGATTTTCATCC 480
QY 168 LeuProAlaProIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
Db 481 CACCCTCTCGAGTCATAACATGGGAGTTCAATCGGACCAACTCTACCTAGTACTATGGAC 540
QY 187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
Db 541 AGGATAACTGCCCTACCAACAGAGTATTGAGATCTATGATGTTCAGCCAAAGGGATTCT 600
QY 207 GlySerTyrArgCysValAlaThrAsnSerAlaArgGlnAspPheSerGlnGluAlaSer 226
Db 601 CGAAATTTATCGTTGATTGCTGCCACTGTAGCCACCGACGTAAGATATGGAGGCTCG 660
QY 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIleVal 246
Db 661 CTAACCTGTGATT-----CCAGCTAAGGAGTCAAAATCTTCCACACACCAATATA 714
QY 247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
Db 715 GCAGGTCCACAGCAACATAACATCTCTTCATCAGACTGTAGTTTGGATGTCATGGCC 774
QY 267 SerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThr 286
Db 775 ACAGGAATCCCAACCAATCATTTCTGGAGCGGCTTGTATCACAATCCATTCATGATGC 834
QY 287 ---AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
Db 835 TTTAATACTCGGGTACTCTGGAATGTTAATCTCATGATATCTGATGTCCAGCTACACAT 894
QY 306 SerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
Db 895 GCTGGAGTATATGTTTGTGGGCCACTACCCCTGGCACACGCAACTTTACAGTTGCTATG 954
QY 326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
Db 955 GCACCTTTAATCTGATTAGTCTCTCTTCAATTTGTTGAATGCCAGAAAGTTTAAACAGG 1014
QY 346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
Db 1015 CCTCGAGCTGGCACTGCTCGATTTGTGTGTCAGGCAGAGGAATCCCTCTCCCAAGATG 1074
QY 366 HisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGlyGly 385
Db 1075 TCATGTTGAAAATAGGAAGGAGATACATTCGATGTGTAGAAATTAATAATGTATAC----- 1128
QY 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTrpGlnCysVal 405
Db 1129 AACAGTAATTTGTTAATTAACAGATATTCTCTGAGATGATGCTATTATTCATGTCATG 1188
QY 406 AlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAlaValValValArgGlu 425
Db 1189 GCTGAGATAGCAAGGATCTATTATTTATCTAGAGCCAGACTGCTGATGATGTCAGAA 1248
QY 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerSerValLeu 445
Db 1249 GACAGACCCAGTCTCCCTTATATGTCATGCTGTAACCATGTCAGAGCTCAGCCATCTT 1308
QY 446 ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisThr 465
Db 1309 TTAGCCTGGAGAGGCGCACTTATTAATTCAGCAAAAGTCATTCCTCCCTTCTGTACTAC 1368
QY 466 GlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAlaAsnAspThrThr 485
```

Db 1369 ATGAAGCAGAGGTTTAAATAATGAAGAGTATCAAGTAGTCATCGGAATGACACAACT 1428
QY 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyr 505
Db 1429 CATTATATTATGATGACTAGAGCTCGCAGCATTTATCTTCTACATTTAGCATAT 1488
QY 506 SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspVal 525
Db 1489 ATGCCAATGGAGCCAGCCAGATGCTGACCATGTGCACAGATACTCTAGAGGATGAC 1548
QY 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
Db 1549 CCCAGA----- 1554
QY 546 TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGluTyr 565
Db 1554 ----- 1554
QY 566 GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
Db 1554 ----- 1554
QY 586 ThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGly 605
Db 1554 ----- 1554
QY 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln 625
Db 1554 ----- 1554
QY 626 SerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuVal 645
Db 1554 ----- 1554
QY 646 ValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrpGly 665
Db 1554 ----- 1554
QY 666 GluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyArgGlyAspGln 685
Db 1554 ----- 1554
QY 686 AlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyrGluLeuThrGln 705
Db 1554 ----- 1554
QY 706 LeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGluAspGly 725
Db 1555 -----AGAAAAATATCATGTGAGACTCTGGCTTACAAACAACATAGACGATGGC 1602
QY 726 TyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuProIleGln 745
Db 1603 TATCAGGCA-----GATCAGACTGTGACGACTCCAGGATGCGTGTCTGTT 1647
QY 746 ArgGly-----ProProLeuProAlaHisValHisAlaGluSerAsnSer 761
Db 1648 CGTGATCGCATGGTCCCTCTCCACCACCCACCACCATCTCTATGCGAAGGCTAACACC 1707
QY 762 SerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsn 781
Db 1708 TCATCTTCCATCTTCTGTCAGCTGAGGAGGCGCTGCATTCACCGCTGCACAAATCATTAAC 1767
QY 782 TyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyrThr 801
Db 1768 TACACCATCGCTGTAAATCTGTGTGGCTGCAGAAATGCTCTTGTGGTCTGTGATCTTCAA 1827
QY 802 SerSerGlyGluAspIleLeuIleGlyLysLeuLysProPheThrLysTyrGluPheAla 821
Db 1828 ACATCAGAAACTCATGTGGTGTGGTCTAGAGTCTAGAACCAACCAAAATACGAATTGCC 1887
QY 822 ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr 841
Db 1888 GTTCGATTACATGCGATCAGCTTTCCAGTCTTGGAGCCCTGTAGTCTACCATTTCTACT 1947

QY 842 LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer 861
Db 1948 CTTCCAGAAACACACAGCAGGCCACCACTGCGAGTAAAGTGACATTAATAGAGGATGAC 2007
QY 862 ThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyrLeu 881
Db 2008 ACTGCCCTGGTCTTCTGGAAAACCCCTGATGGCCAGAAACAGTGTGTGACCGCTATACT 2067
QY 882 IleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThrGlu 901
Db 2068 ATCTTATATGATCATCTAGGAAGCCCTGGATTGCGAGGAGTGGCAGGTCTTACACCGTAA 2127
QY 902 GlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLys 921
Db 2128 GGGGCAATAACCATGGCTTTGCTAGAAAACTTGGTAGCAGGAAATGTGTACATTGTCAAG 2187
QY 922 MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
Db 2188 ATATCTGCATCCAATGAGGTGGAGAGGCCCTTTTCAAAATTTCTGTGGAGCTGGCAGTA 2247
QY 942 LeuGlnGluThrPheSerAspSer----- 949
Db 2248 CTTCCAAAGGAAACCTCTGAATCAATCAGAGGCCCAAGCGTTTAGATTCTGCTGATGCC 2307
QY 950 -----LeuAspValHisAlaValThrGlyIleIleValGly 961
Db 2308 AAAGTTTATTCAGGATATTACCATTCTGCACCAAAAAATCAATGACTGGCATTTGCTAGGT 2367
QY 962 ValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSerSer 981
Db 2368 GTTGGCATAGCCTTGACCTGCATCTCATCTGTGTCTCATCTTGTATATACCGAAGTAA 2427
QY 982 HisArgGluAla-----LeuProGlyLeu 989
Db 2428 GCCAGAAATCATCTGTTCCAGACGGCAGCAAGATGGAACTCAACAGTTACTCTGTACC 2487
QY 990 SerSerSerGlyThrProGlyAsn 997
Db 2488 AGTGCCTCTCTAGTGTGAAAT 2511

RESULT 15

US-10-231-353-15
; Sequence 15, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NO. US20030023064A1e1 Human Phosphatases and Polynucleotides Et.
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231,353
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-231-353-15

Alignment Scores:
Pred. No.: 8,66e-98 Length: 2481
Score: 1329.00 Matches: 312
Percent Similarity: 46.29% Conservative: 131
Best Local Similarity: 32.60% Mismatches: 276
Query Match: 20.12% Indels: 238

DB:	14	Gaps:	11
US-09-754-997A-2 (1-1252) x US-10-231-353-15 (1-2481)			
QY	79	GluHisGluAAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlu	98
DB	7	GAATAAAGCGATGAGGCTCTTCTAACGGCTCTTTATACATCAGT-----GAG	57
QY	99	GlnGluAspSerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTyr	118
DB	58	GTGGAAGCGAGCGAGGAGCGAGTCC-----GATGAAGGATTTTAT	99
QY	119	SerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeu	138
DB	100	CAGTGTCTGGCAATGAACAATATGAGCCATCTTAGTCAAAAGCTCATCTCTGCTTA	159
QY	139	AlaThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThr	158
DB	160	TCAACTATTCTGCAATTTGAGTCCAGCCCAATTTCCACTGAGTCCACGAAGTGGAGTT	219
QY	159	AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAsp	178
DB	220	GTCGATTTGCAATGCAAGATTTCATCCGACCTCTCTGCGAGTCATTAACATGGAGTTCAAT	279
QY	179	GlnValThrValPro---GluGluProArgLeuIleThrLeuProLysTrpLeuLeuGln	197
DB	280	CGGACAACTCTACCTATGACTATGACAGGATAAATGCGCTTACCAACAGGAGTATGCGAG	339
QY	198	IleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAla	217
DB	340	ATCTATGATGTGCGCAAGGATTTCTGGAATATTCGTTGATTTGCTGCCACTGTAGCC	399
QY	218	ArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAla	237
DB	400	CACGACGTAAGATGATGAGCGCTCGTAACTGTGATT-----CCAGCTAAGGAGTCA	453
QY	238	ThrArgGlyGlnAspValValIleValAlaAlaProGluAsnThrThrValValSerGly	257
DB	454	AAATCCTTCCACACACCAACCAATATAGCAGTCCACAGAACATAACAACATCTTCAT	513
QY	258	GlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpVal	277
DB	514	CAGACTGTAGTTTGGAAATGATGCGCACAGGAAATCCCAACCAATCATTTCTGGAGC	573
QY	278	ArgGlnAspGlyProLysSerThr---AspValIleValLeuGlyArgThrAsnLeu	296
DB	574	CGCCTTGATCACAAATCCATTCATGTCTTTAATATCTCGGGTACTTGGAAATGGTAACTCTC	633
QY	297	LeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysPro	316
DB	634	ATGATATCTGTATGCTCAGGTACACATGCTGGAGTATATGTTTGTGGGCCCACTACCCCT	693
QY	317	LeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle	336
DB	694	GGCACAGCAACTTTACAGTGTCTATGGCAACTTTAACTGTATTAGTCTCTCTTCATTT	753
QY	337	SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaAlaPheValCysArg	356
DB	754	GTGGAATGGCCAGAAAGTTTAAACAGGCGCTCGAGCTGCGACTGCGATTTGTGTGTCAG	813
QY	357	AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgPro	376
DB	814	GCAGAGGAATCCCTCTCCAGATGTCATGTTGAAAATGGGAAGGAGATACATTCG	873
QY	377	AsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeu	396
DB	874	AATGGTAGAATTTAAATGTAC-----AACAGTAAATTTGTAATTAACAGATTATTCCT	927
QY	397	GlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAla	416
DB	928	GAAGATGATGCTATTATCATGTGCTGAGTGAATAGCAAGGATCTATTTTATCTAGA	987
QY	417	AlaProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThrAla	436
DB			
DB	988	GCCAGACTGACTGTAGTGAATGTCAAGACAGACCCAGTGTCTCCCTATAATAATGTACATCT	1047
QY	437	ThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlu	456
DB	1048	GAACACCATGTCAAGCTCAGCCATCTTTTACCTCGGAGAGGCCACTTTTAAATTCAGAC	1107
QY	457	GlnIleIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyr	476
DB	1108	AAAGTCATTGCTATTCTGTACACTACATGAAGACAGAGGCTTTAAATAATGAAGAGTAT	1167
QY	477	GlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr	496
DB	1168	CAAGTAGTCATCGGAATATGACCAACTCATTTATTTATTTGATGATCTTAGAGCTGCCAGC	1227
QY	497	AspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerPro	516
DB	1228	AATTATATCTTCTACATTTGATGATATATGCAATGGGAGCCAGCATGTCTGACCAT	1287
QY	517	AlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSer	536
DB	1288	GTGACACAGATACTCTAGAGGATGACCCAGA-----	1320
QY	537	ProAsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGly	556
DB	1320	-----	1320
QY	557	GlnValLeuLysTyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThr	576
DB	1320	-----	1320
QY	577	GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr	596
DB	1320	-----	1320
QY	597	ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln	616
DB	1320	-----	1320
QY	617	HisArgThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLys	636
DB	1320	-----	1320
QY	637	ValArgAlaLysMetGluSerLeuValValSerTrpGlnProProProHisProThrGln	656
DB	1320	-----	1320
QY	657	IleSerGlyTyrLysLeuTyrTrpGlyGluValGlyThrGluGluAlaAspGlyAsp	676
DB	1320	-----	1320
QY	677	ArgProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLys	696
DB	1320	-----	1320
QY	697	LysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeu	716
DB	1321	-----AGAAAATATCATGTGAGACTC	1341
QY	717	ValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLys	736
DB	1342	CTGGCTTACACACACATAGACGATGGCTATCAGGCA-----GATCAGACT	1386
QY	737	AlaProThrProAspLeuProIleGlnArgGly-----ProProLeuProPro	752
DB	1387	GTCCAGCACTCCAGGATGCGTGTCTGTTGATGTCATGTCATGTCCTCCACACACCC	1446
QY	753	AlaHisValHisAlaGluSerAsnSerSerThrSerIleTrpLeuArgTrpLysLysPro	772
DB	1447	CACCATCTCTATGCGAAGGCTAACACCTCATCTTCCATCTTCTGCACTGGAGGAGGCT	1506
QY	773	AspPheThrThrValLysIleValAsnTyrThrValArgPheGlyProTyrGlyLeuArg	792

```
Db 1507 GCATTCACCGCTGCACAAATCAATTAACATACCATCCGCTGTAATCTCTGTGGCCTGCAG 1566
Qy 793 AsnAlaSerLeuValThrTyrTyrThrSerSerGluAspIleLeuIleGlyGlyLeu 812
Db 1567 AATGCTTCTTTGGTTCTGTACCTTCAACATCAGAACTCAGATGTGGTTCAGGTCTA 1626
Qy 813 LysProPheThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyPro 832
Db 1627 GAACCAACACACCAAAATACCAATTTGCGGTTTCGATTACATGTGGATCAGCTTCCAGTCCT 1686
Qy 833 PheGlySerValValGluArgSerThrLeuProAspArgProSerThrProProSerAsp 852
Db 1687 TGGAGCCCTGTAGTCTACCATCTCTTCCAGAACACACAGCAGCCACCCAGTTGGA 1746
Qy 853 LeuArgLeuSerProLeuThrProSerThrValArgLeuHisTyrCysProProThrGlu 872
Db 1747 GTAAAGTAGACATTATAGAGGATGACATGCCCTGGTTCTTGGAAACCCCTGATGGC 1806
Qy 873 ProAsnGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGlu 892
Db 1807 CCAGAAACAGTGTGACCCGCTATCTATCTTATATGCATCTAGGAAGGCTGGATTGCA 1866
Qy 893 HisGlnTyrThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeu 912
Db 1867 GGAGAGTGGCAGGTCTTACACCGTGAAGGGGCAATAACCATGGCTTTGCTAGAAAACCTG 1926
Qy 913 GluSerAspThrArgTyrPhePheLeuMetGlyAlaArgThrGluValGluValGlyPro 932
Db 1927 GTAGCAGGAATGTGTACATTGTCAAGATATCTGCATCCATGAGGTGGAGAGGACCC 1986
Qy 933 PheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSer----- 949
Db 1987 TTTTCAAATTCCTGGAGCTGGCAGTACTTCCAAAGGAAACCTCTGAATCAAATCAGAGG 2046
Qy 950 -----LeuAspVal 952
Db 2047 CCCAAGCGTTTAGATTCTGCTGATGCCAAAGTTTATTTCAGGATATTACCATCTGGACCAA 2106
Qy 953 HisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuAlaCys 972
Db 2107 AATCAATGACTGGCATTTGCTGTAGGTGTGGCATGACCTTGACCTGCATCCTCATCTGT 2166
Qy 973 MetCysAlaGlyLeuArgGlnSerSerHisArgGluAla----- 985
Db 2167 GTTCTCATCTTGATATACCAAGTAAAGCCAGGAATCATCTGCTTCCAGACGGCACAG 2226
Qy 986 -----LeuProGlyLeuSerSerSerGlyThrProGlyAsn 997
Db 2227 AATGGAACCTCAACAGTTACCTCGTCCAGTGCCTCTTGTAGTAGTGAAT 2277
```

Search completed: October 14, 2003, 09:45:00

Job time : 1026 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 14, 2003, 06:11:10 ; Search time 137 Seconds
(without alignments)
4033.662 Million cell updates/sec

Title: US-09-754-997A-2
Perfect score: 6504
Sequence: 1 MARADTGRLLVLTFCLLSA.....CPLTVSPSLPRAPVSSAQVP 1252

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09754997/runat_14102003_061101_9736/app_query.fasta_1.1415
-DB=Issued_Patents_NA -QMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09754997 SCGN 1.1.52 @runat_14102003_061101_9736 -NCPV=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1747.5	26.5	3453	US-09-877-730-7
2	1746.5	26.4	3210	US-09-877-730-1
3	1746.5	26.4	3874	US-09-877-730-31
4	1691.5	25.6	2976	US-09-877-730-11
5	1691.5	25.6	3219	US-09-877-730-17
6	1384	21.0	2715	US-09-877-730-5
7	1384	21.0	2958	US-09-877-730-9
8	1329	20.1	2481	US-09-877-730-15
9	1329	20.1	2724	US-09-877-730-19
10	1144	17.3	2139	US-09-877-730-21
11	1144	17.3	2382	US-09-877-730-27
12	943	14.3	4608	US-09-041-886-24

13	943	14.3	4608	5	PCT-US94-05277-1	Sequence 1, Appli
14	909	13.8	1875	4	US-09-877-730-23	Sequence 23, Appl
15	781.5	11.8	1644	4	US-09-877-730-25	Sequence 25, Appl
16	781.5	11.8	1887	4	US-09-877-730-29	Sequence 29, Appl
17	647.5	9.8	1143	4	US-09-877-730-3	Sequence 3, Appli
18	638	9.7	6000	1	US-08-348-006B-6	Sequence 6, Appli
19	638	9.7	6000	2	US-08-800-825A-6	Sequence 6, Appli
20	638	9.7	6000	3	US-09-158-657-6	Sequence 6, Appli
21	638	9.7	6000	5	PCT-US94-10166-6	Sequence 6, Appli
22	634	9.6	4078	4	US-09-016-434-1132	Sequence 1132, Ap
23	592.5	9.0	909	4	US-09-877-730-13	Sequence 13, Appl
24	548.5	8.3	3888	4	US-08-506-296B-13	Sequence 13, Appl
25	540.5	8.2	3943	4	US-08-506-296B-27	Sequence 27, Appl
26	533	8.1	3774	2	US-08-341-843B-1	Sequence 1, Appli
27	533	8.1	3774	2	US-08-427-497B-2	Sequence 1, Appli
28	533	8.1	3774	2	US-08-427-497B-2	Sequence 2, Appli
29	522	7.9	5690	2	US-08-447-464-2	Sequence 2, Appli
30	522	7.9	5690	2	US-08-716-679-2	Sequence 2, Appli
31	514	7.8	3783	4	US-08-506-296B-20	Sequence 20, Appl
32	500.5	7.6	3189	2	US-08-427-497B-3	Sequence 3, Appli
33	500.5	7.6	5824	4	US-09-620-312D-72	Sequence 72, Appl
34	477	7.2	3991	4	US-08-506-296B-3	Sequence 3, Appli
35	443	6.7	6814	4	US-09-484-970B-66	Sequence 66, Appl
36	436.5	6.6	3360	1	US-08-408-093-5	Sequence 5, Appli
37	436.5	6.6	3360	1	US-08-408-420A-5	Sequence 5, Appli
38	436.5	6.6	3360	1	US-08-714-901-5	Sequence 5, Appli
39	436.5	6.6	3360	3	US-08-040-741-5	Sequence 4, Appli
40	387	5.9	2600	2	US-08-427-497B-4	Sequence 4, Appli
41	349.5	5.3	13857	4	US-09-620-312D-75	Sequence 75, Appl
42	324.5	4.9	3024	1	US-07-923-376-7	Sequence 7, Appli
43	320	4.8	4078	4	US-09-016-434-1120	Sequence 1120, Ap
44	320	4.8	4724	1	US-08-404-665-3	Sequence 3, Appli
45	320	4.8	4724	1	US-08-404-671-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-877-730-7
; Sequence 7, Application US/098777730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632e1 Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3453
; TYPE: DNA
; ORGANISM: homo sapiens

Alignment Scores:
Pred. No.: 4.42e-106 Length: 3453
Score: 1747.50 Matches: 397
Percent Similarity: 54.95% Conservative: 169
Best Local Similarity: 38.54% Mismatches: 381
Query Match: 26.46% Indels: 83
DB: 4 Gaps: 18

US-09-754-997A-2 (1-1252) x US-09-877-730-7 (1-3453)

QY 9 GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln 28
| | | | |
Db 43 GGGATGCTGCTCGCGCGCTCCTGCTCCTGCTGCTCAGT-----CCTTGGCCAGA 96
| | | | |
QY 29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValleLeuGlyPro 47
| | | | |
Db 97 GTGTGTGCTTTAGCGAAGCTGCTTTTGTAAAGAACACACAGGATGTAACGTGTCAAGA 156
| | | | |
QY 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArg 67
| | | | |
Db 157 AAGACCCAGTCTGTTTAGATCCAGGCTCACGGAGAAGTT-----CCTATTAAAG 207
| | | | |
QY 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro 87
| | | | |
Db 208 GTCACATGTTGAAATGGAGCAAAATGCTCGAAATAAACGGATCGAGGTCTTCTTCT 267
| | | | |
QY 88 AsnGlySerLeuTrpLeuSerProLeuGluGlnGluAspSerAspAspGluGluAla 107
| | | | |
Db 268 AACGGCTCTTTATACATCACT-----GAGGTGGAAGGCAGCGGAGAGCAGTC 318
| | | | |
QY 108 LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly 127
| | | | |
Db 319 -----GATGAAGGATTTTATCAGTCTTGGCAATGAACAAATATGGA 360
| | | | |
QY 128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
| | | | |
Db 361 GCCATCTTAGTCAAAAGCTCATCTGCTTATCACTATTCTGCAATTTGAAGTCCAG 420
| | | | |
QY 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
| | | | |
Db 421 CCAATTCCACTGAGGTCCAGGAAGGTGGAGTTGCTCGATTTCATGCAAGATTTTCATCC 480
| | | | |
QY 168 LeuProAlaProIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
| | | | |
Db 481 CACCCTCTCGATCATAAATGGAGTTTCAATCGGACAACTCTACTATGACTATGGAC 540
| | | | |
QY 187 ArgLeuIleThrLeuProLysTrpLeuGlnIleLeuAspValGlnAspSerAspAla 206
| | | | |
Db 541 AGGATAACTGCCCTACACAGAGATGATTCAGATCTATGATGTCAGCCAAAGGATCT 600
| | | | |
QY 207 GlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnIleAlaSer 226
| | | | |
Db 601 GGAATATATCGTTGTTATGCTGCCACTGTAGCCACCGAGCTAAAGTATGGAGGCTCG 660
| | | | |
QY 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIleVal 246
| | | | |
Db 661 CTAACGTGTGAT-----CCAGCTAAGGAGTCAAAATCTTCCACACACCAACAATTATA 714
| | | | |
QY 247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
| | | | |
Db 715 GCAGGTCCAGAACATACACATCTTTCATCAGACTGTAGTTTGGAAATGCATGGCC 774
| | | | |
QY 267 SerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThr 286
| | | | |
Db 775 ACAGGAATCCCAAAACCAATCATTTCTTGAGCGCGCTTCATCACAAATCCATTGATGC 834
| | | | |
QY 287 ---AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
| | | | |
Db 835 TTTAATACTCGGGTACTTGGAAATGGTAATCTCATGTATCTCATGTATCTGATGCTACACAT 894
| | | | |
QY 306 SerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
| | | | |
Db 895 CTGGAGTATGTTTGTGGGCCCATACCCCTGGCACGCAACTTACAGTTGCTGCTATG 954
| | | | |
QY 326 AlaGluLeuArgValLeuAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
| | | | |
Db 955 GCACTTTAACTGTATTAGTCTCTCTTCTTATTGTTGAATGGCCAGAAAGTTTAAACAAG 1014
| | | | |
QY 346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
| | | | |
Db 1015 CCTCGAGCTGGCACTGCTCGATTTGTGTGTCAGCAGGAAGAAATCCCTCTCCCAAGATG 1074
| | | | |
QY 366 HistTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGlyGly 385
| | | | |

Db 1075 TCATGTTGAAAAATGGGAAGAGATACATTCGAATGTTAGAAATTAATAATGTAC----- 1128
| | | | |
QY 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysVal 405
| | | | |
Db 1129 AACAGTAATTTGGTAATTAACCGAGATATTCTCTGAAGATGATGCTATTATTATCAGTGCATG 1188
| | | | |
QY 406 AlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValValValArgGlu 425
| | | | |
Db 1189 GCTGAGATAGCCAAAGGATCTATTATTATCTAGAGCCAGACTGACTGTAGTGTAGTCAGAA 1248
| | | | |
QY 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerSerValLeu 445
| | | | |
Db 1249 GACAGACCAGTGTCCCTATATAATGTACATGCTGAAACCATGTCAAGCTCAGCACTTCTT 1308
| | | | |
QY 446 ValAlaTrpGluAspProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr 465
| | | | |
Db 1309 TTAGCCTGGAGAGGCCACTTTTATAATTCAGACAAAGTCATTGCTATTCTGTACACTAC 1368
| | | | |
QY 466 GlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThr 485
| | | | |
Db 1369 ATGAAGCAGAGGTTTAAATTAATGAAGATGATCAAGTAGTCAATCGGAATGACCAACT 1428
| | | | |
QY 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyr 505
| | | | |
Db 1429 CATTTATATTATGATGACTTAGAGCTGCAGCAATTTATCTTCTCATTTGTAGCATAT 1488
| | | | |
QY 506 SerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHisThrLeuAspAspVal 525
| | | | |
Db 1489 ATGCCAATGGGAGCCAGCCAGATGCTCACCATTGTCAGACACAGATACTCTAGAGATGTT 1548
| | | | |
QY 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
| | | | |
Db 1549 CCCCTGAGACCTCTCGAAATTTAGTTTGACAGTCCAGTCCACATGATTTCTCATCTCC 1608
| | | | |
QY 546 TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGluTyr 565
| | | | |
Db 1609 TGGTGCCTCAATCCAGCCAAATATCGGGGGCCAAAGTGGTGTCTGTATCGTGTCTTTC 1668
| | | | |
QY 566 GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
| | | | |
Db 1669 CCCCCTAAGTACTGAGAAATTCATCAATCCAGTCTCTGAGTCCCGGGGACCCAGCATGATAC 1728
| | | | |
QY 586 ThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGly 605
| | | | |
Db 1729 CTTTGTGAAGGCTGAACCTGACAGTGTCTACCTGTTCCGATTTACTCTGCCACCAGA 1788
| | | | |
QY 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln 625
| | | | |
Db 1789 GTGGGGCTGGGAGAGTCATCAGTATGGACTTCACATAGGACGCCCAAGCT---ACAAGC 1845
| | | | |
QY 626 SerHisValProPheAlaProAlaGluLeuLysValArg---AlaLysMetGluSerLeu 644
| | | | |
Db 1846 GTGAAGGCCCTTAAGTCTCCA---GAGTTGCAATTTGGAGCCTCTGAACTGTATCCACCAT 1902
| | | | |
QY 645 ValValSerTrpGlnProProHisProThr---GlnIleSerGlyTyrLysValLeuTyr 663
| | | | |
Db 1903 TCTGTGAGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTTCAGGGCTACAGCTGTATC 1962
| | | | |
QY 664 TrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProGlyGlyArgGly 683
| | | | |
Db 1963 TACAAGGAAGAGGCGCAGCGAGGAAT----- 1989
| | | | |
QY 684 AspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysValLysGlnTyrGluLeu 703
| | | | |
Db 1990 -----GGGCCCATTTTCTTGGATACCAAGGACCTACTCTATATCTCTC 2031
| | | | |
QY 704 ThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGlu 723
| | | | |
Db 2032 AGTGGCTTAGACCCCAAGAAATATCATGTGAGACTCCTGGCTTACAAACACATAGAC 2091
| | | | |
QY 724 AspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuPro 743
| | | | |

```
Db 2092 GATGCTATCAGGCA-----GATCAGACTGTCCAGCACTCCAGGATGCGTG 2136
Qy 744 lIeGlnAArgGly-----ProLeuPProAlaHisValHisAlaGluSer 759
Db 2137 TCTGTGTCGATGCGATGCTCTCTCCACCAACCCACCACTCTCTATGCGAGGCT 2196
Qy 760 AnSerSerThrSerIleThrLeuArgTrpLysLysProAspPheThrThrValLysIle 779
Db 2197 AACACCTCATCTTCCATCTCTCTGCACTGGAGAGGCTGCACTTCCACGCTGCACAAATC 2256
Qy 780 ValAenThrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTrp 799
Db 2257 ATTAACTACACCATCGCTGTAACTCTGTGCGCTGCAGAAATGCTCTTGTGTTCTGTAC 2316
Qy 800 TyrThrSerSerGlyGluAspIleLeuIleGlyGlyLeuLysProPheThrLysTrpGlu 819
Db 2317 CTTCAACATCAGAACTACATGTTGTTCAAGTCTAGAACCAACCAACCAATACGAA 2376
Qy 820 PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg 839
Db 2377 TTTGCGGTTTCGATTACATCTGCATCGATCTTCCAGTCTTCCAGTCTTGGAGCCCTGTAGCTACCAT 2436
Qy 840 SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr 859
Db 2437 TCTACTCTTCCAGAACGACACAGAGGCCACCACTGTTGGAGTAAAGTGAATATATAGAG 2496
Qy 860 ProSerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluValGlu 879
Db 2497 GATGACACTCGCTGGTCTTCTGGAACCCCTGTATGGCCCAAGAACAGTGTGACCCGC 2556
Qy 880 TyrLeuIleLeuTrpSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThr 899
Db 2557 TATACTATCTTATGATCATCTAGGAGGCTGAGTGTGAGGAGTGGCAGGCTCTTACAC 2616
Qy 900 ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTrpPhe 919
Db 2617 CTGTAAGGGGCAATACCACTGCTTGTAGAAAACCTTGTAGCAGGAAATGTGTACATT 2676
Qy 920 PheLysMetGlyValAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
Db 2677 GTCAGATATCTGCATCCATGAGTGGAGGAGGAGGACCCCTTTTCAATCTCTGGAGCTG 2736
Qy 940 IleThrLeuGlnGluThrPheSerAspSer-----LeuAspValHisAlaValThrGlyIleIle 949
Db 2737 GCAGTACTTCCAAAGGAACCTCTGAATCAATCAATCAGAGCCCAAGCGTTTAGATTCTGCT 2796
Qy 950 -----LeuAspValHisAlaValThrGlyIleIle 959
Db 2797 GATGCCAAAGTTTATTTCAGGATATTACCATCTGGACCAAAAATCAATGACTGGCATTTGCT 2856
Qy 960 ValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln 979
Db 2857 GTAGGTGTTGGCATAGCCCTTGACCTGCATCTCTCATCTGTGTTCTCATCTTTGATATACCGA 2916
Qy 980 SerSerHisArgGluAla-----LeuPro 987
Db 2917 AGTAAAGCCAGGAATCATCTGCTTCCAGACGGCACAGATGGAAGTCAACAGTTACTCT 2976
Qy 988 GlyLeuSerSerSerGlyThrProGlyAsn 997
Db 2977 CGTACCAGTGCCTCTTAGCTAGTGGAAAT 3006
```

RESULT 2

```
US-09-877-730-1
; Sequence 1, Application US/0987730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
```

```
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632e1 Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3210
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-1
Alignment Scores:
Pred. No.: 4,66e-106 Length: 3210
Score: 1746.50 Matches: 397
Percent Similarity: 54.95% Conservative: 169
Best Local Similarity: 38.54% Mismatches: 381
Query Match: 26.45% Indels: 83
DB: 4 Gaps: 18
```

US-09-754-997A-2 (1-1252) x US-09-877-730-1 (1-3210)

```
Qy 9 GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGlyLeuProLeuProGln 28
Db 43 GGGATGCTGCTCCGCGGCTCTCTGCTCTGCTGCTCAGT-----CCTTTGCCAGGA 96
Qy 29 GluThrThrVal-----LysLeuSerCysAspGluGlyProLeuGlnValIleLeuGlyPro 47
Db 97 GTGTGTGCTTTTAGCGAACTGCTCTTTGTAAGAAGAACCAACAGGATGTAACTGTCAAGA 156
Qy 48 GluGlnAlaValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArg 67
Db 157 AAGGACCAGTCGTTTAGATTGTCAGGCTCAGGAGAAAGTT-----CCTATTAAAG 207
Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro 87
Db 208 GTCACATGTTGAAAAATGGAGCAAAATGCTGAAATAAAACGATCGAGGTTCTTTCT 267
Qy 88 AsnGlySerLeuTrpLeuSerSerProLeuGlnGluAspSerAspAspGluAla 107
Db 268 AACGGCTCTTTATACATCAGT-----GAGGTGGAAGGCGAGGAGGAGCAGTCC 318
Qy 108 LeuArgIleThrLysValThrGluGlySerTrpSerCysLeuAlaHisSerProLeuGly 127
Db 319 -----GATGAAGGATTTTATCAGTCTCTGGCAATGAACAAATATGGA 360
Qy 128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
Db 361 GCCATTCTTAGTCAAAAAGCTCATCTTGCCTTATCACTATTCTTGCATTGGAAGTCCAG 420
Qy 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
Db 421 CCAATTTCCACTGAGGTCACAGAGTGGAGTGTCTGATTTGCATGCAAGATTTTCATCC 480
Qy 168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro-----GluGluPro 186
Db 481 CACCCTCTGCGATCATCAACATGGAGTTCAATCGACACACTCTACCTATGATATGGAC 540
Qy 187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
Db 541 AGGATACTGCCCTTACCACAGGAGTATTGCGATCTATGATGTGAGCCAAAGGATTTCT 600
Qy 207 GlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArgPheSerSerGlnGluAlaSer 226
Db 601 GGAATATTATGTTGTTGTTGCTGCCACTGTAGCCACCGACGCTAAAGATATGGAGCTCG 660
Qy 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValIleVal 246
Db 661 CTAACCTGTGATT-----CCAGCTAAGGAGTCAAAATCCTTCCACACACCAATTTATA 714
```

Qy	247	AlaAlaProGluAenThrThrValValSerGlyGlnAsnValValMetGluCysValAla	266
Db	715	GCAGGTCCACAGAACATAAACACATCTCTTCATCAGACTGTAGTGTGGAATGATGCACGCC	774
Qy	267	SerAlaAspProThrProPheValSerTrpValArgGlnAspGlyIysProIleSerThr	286
Db	775	ACAGGAATCCCAACCAACATCATTTCTTGAGCGCCCTTGATCACAANTCATGATGTC	834
Qy	287	--- AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis	305
Db	835	TTTAATACTCGGGTACTTGGAAATGGTAATCTCATGATATCTGATGTCAGGCTACACAT	894
Qy	306	SerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla	325
Db	895	GCTGGAGTATATGTTGTGCGGCCACTACCCCTGGCACAGCAACTTTACAGTTGCTATG	954
Qy	326	AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg	345
Db	955	GCAACTTTAACTGATTAGTCTCTCTTCATTTGTAATGGCCAGAAAGTTTAACAAGG	1014
Qy	346	ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu	365
Db	1015	CCTCGAGCTGGCACTGCTGATTTGTGTGTCAGGCAGAGGAATCCCTCTCCCAAGATG	1074
Qy	366	HisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValIysValGlnGly	385
Db	1075	TCATGTTGMAAAATCGAAGGAAGATACATTCGAATGGTAGAATTAATAATGTAC----	1128
Qy	386	GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysVal	405
Db	1129	AACAGTAAATTTGGTAATTAACCAATATTCTCTGAAGATGATGCTATTTATCAGTGCATG	1188
Qy	406	AlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgGlu	425
Db	1189	GCTGAGAAATAGCCAGGATCTATTTATCTAGACCCAGACTGACTGTAGTGATGTACGAA	1248
Qy	426	GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu	445
Db	1249	GACAGACCCAGTGTCTCCCTATAATGTATACATGCTGAAACCATGTCAAGCTCAGCCATCTT	1308
Qy	446	ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr	465
Db	1309	TTAGCTCTGGAGAGCCACTTTATTAATTCAGACAAAGTCATGCTCTATTCTGTACACTAC	1368
Qy	466	GlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThr	485
Db	1369	ATGAACAGCAAGAGTTTAATTAATGAGAGTATCAGTAGTATCAGGAATATGACCAACT	1428
Qy	486	GluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyr	505
Db	1429	CATTATATTATGATGACTTAGAGCTGCCAGCAATTATATCTTCTACATTGTAGCATAT	1488
Qy	506	SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAspVal	525
Db	1489	ATGCCAATGGGAGCCAGCCAGATGCTCTGACCATGTGACACAGAATATCTCTAGAGGATGT	1548
Qy	526	ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla	545
Db	1549	CCCTTGAGACCTCTCGAAATTAGTTTGACAAAGTCGAAGTCCCATGTATTTCTCATCTCC	1608
Qy	546	TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGluTyr	565
Db	1609	TGGCTGCCAATCCCGACCAATATCGCGGGGCCCAAGTGGTGCTGATCGCTGCTCTTTC	1668
Qy	566	GlyLeuGlyIysGluAspGlnValPheSerThrGluValProGlyIysAsnGluThrGlnLeu	585
Db	1669	CGCCTAAGTACTGAGAATTCAATCCAAGTTCTGAGAGTCTCCCGGGGACCAACGATGATAC	1728
Qy	586	ThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGly	605
Db	1729	CTTTTGGAGGCCCTGAACACCTGACAGTGTCTACTGTTGCGATTAATCTGCTGCCACAGA	1788
Qy	606	AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln	625

Db	1789	GTGGGGCTGGAGAGTCATCAGTATGGACTTCATCAGACGCCAAAGCT---ACAAGC	1845
Qy	626	SerHisValProPheAlaProAlaGluLeuLysValArg---AlaLysMetGluSerLeu	644
Db	1846	GTGAAAGCCCTAAGTCTCCA---GAGTTGCATTGGAGCCTCTGAACTGTACCACCAATT	1902
Qy	645	ValValSerTrpGlnProProHisLysProThr---GlnIleSerGlyTyrIysLeuTyr	663
Db	1903	TCTGTAGGTGGCAGCAGATGTPAGAGGACACACAGCTGTCTATTTCAGGGCTACACAGCTGTATC	1962
Qy	664	TrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyArgGly	683
Db	1963	TACAAGGAAGAGGGCAGCAGGAGAAAT-----	1989
Qy	684	AspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyrGluLeu	703
Db	1990	-----GGGCCCATTTCTGGATACCAAGGACCTACTCTATATCTCTC	2031
Qy	704	ThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGlu	723
Db	2032	AGTGGCTTAGACCCCAAGAGAAATATCATGTGAGACTCTGTGCTTACCAACACATAGAC	2091
Qy	724	AspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuPro	743
Db	2092	GATGGCTATCAGCA-----GATCAGACTGTTCAGCACTCCAGGATCGGTG	2136
Qy	744	IleGlnArgGly-----ProProLeuProProAlaHisValHisAlaGluSer	759
Db	2137	TCTGTTGGTATCGCATGGTCCCTCTCTCCACCACCCACCACCATCTCTATGGCAAGCT	2196
Qy	760	AsnSerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIle	779
Db	2197	AACACCTCATCTCCATCTTCTCTGCATGGAGAGGCGTGCATTTCCCGCTGCACAAATC	2256
Qy	780	ValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyr	799
Db	2257	ATTAACATACACCATCCGCTGTAAATCTCTGTGGCTCGCAGAATGCTCTTTGGTCTGTATC	2316
Qy	800	TyrThrSerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrLysTyrGlu	819
Db	2317	CTTCAACATCAGAAATCAGTGTGGTTGGTTCAAGGTCTAGAACCAACCAATATAGAA	2376
Qy	820	PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg	839
Db	2377	TTTGCCGTTTCGATTACATGGATCAGCTTTCCAGTCTTGGAGCCCTGTAGTCTACCAT	2436
Qy	840	SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr	859
Db	2437	TCTACTCTTCAGACGACCAGCAGGCCACCACTGTGGAGTAAAGTGCATTAATAGAG	2496
Qy	860	ProSerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGlu	879
Db	2497	GATGACACTGCCCTGGTTTCTTGAAACCCCTGTATGGCCCAAGACAGTTGTGACCCGC	2556
Qy	880	TyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThr	899
Db	2557	TATACTATCTTATATGTCATCTAGAAAGCCCTGGATTTCAGAGAGTGGCAGGCTTTACAC	2616
Qy	900	ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhe	919
Db	2617	CGTGAAGGGGCAATAACCATGGCTTTGCTAGAAAACTTGTGTAGCAGGAAATGTGTACATT	2676
Qy	920	PheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal	939
Db	2677	GTCAGATATCTGTCATCCATAGAGGTGGGAGAAGGACCCCTTTTCAAATTTCTGTGGAGCTG	2736
Qy	940	IleThrLeuGlnGluThrPheSerAspSer-----	949
Db	2737	GCAGTACTTCCAAAGGAAACCTCTGATCAATCATCAGAGGCCCAAGCGTTTAGATTCTGCT	2796
Qy	950	-----LeuAspValHisAlaValThrGlyIleIle	959

```

Db      2797 GATGCCAAAGTTTATTCAGATATTTACCATCTGGACCAAAAATCAATGACTGGCATGCT 2856
Qy      960 ValGlyValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGln 979
      2857 GTAGGTGTTGGCATACCTTGACCTGCATCTCTCATCTGTGTTCTCATCTTGATATACCGA 2916
Qy      980 SerSerHisArgGluAla-----LeuPro 987
      2917 AGTAAAGCCAGGAATCATCTGCTTCCAGACGGCACAGATGGAATCAACAGTTTACCT 2976
Qy      988 GlyLeuSerSerGlyThrProGlyAsn 997
      2977 CGTACAGTGCCTCTTACCTAGTGGAAAT 3006

RESULT 3
US-09-877-730-31
; Sequence 31, Application US/09877730
; Patent No. 645632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 645632el Human Phosphatases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 3874
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-31

Alignment Scores:
Pred. No.:      6,048-106      Length:      3874
Score:          1746.50      Matches:      397
Percent Similarity: 54.95%      Conservative: 169
Best Local Similarity: 38.54%      Mismatches:  381
Query Match:      26.45%      Indels:       83
DB:               4          Gaps:         18

US-09-754-997A-2 (1-1252) x US-09-877-730-31 (1-3874)
Qy      9 GlyLeuLeuValThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln 28
      233 GGGATGCTGCTCGCGCGCTCTGCTCTGCTGCTGCTCAGT-----CCTTTGCCAGA 286
Qy      29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValIleLeuGlyPro 47
      287 GTGTGTGCTGTTAGCGACTGCTTTTGTGAAAAGAACACACAGATGTTAACTGTCAAGA 346
Qy      48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProThrArg 67
      347 AAGGACCCAGTCGTTTACATTTGCCAGGCTCACGGAGAGTT-----CCTATTAA 397
Qy      68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuPro 87
      398 GTCATAGTTGGAATAATGGACAAAATGCTCGAAAATAAACCAGGATCGAGGTTCTTCT 457
Qy      88 AsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAspSerAspAspGluGluAla 107
      458 AACGGCTCTTTATACATCAGT-----GAGGTGGAAGGCGGAGGAGGAGGAGTCC 508
Qy      108 LeuArgIleTrpLysValThrGluGlySerThrCysLeuAlaHisSerProLeuGly 127
      509 -----GATGAAGGATTTTATCATGTCGTCTGGCAATGAACAAATATGGA 550

```

```

Qy      128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
      551 GCCATTCTTAGTCAAAAAGCTCATCTGCTTATCACTATTCTTGTGCAATTTGAAGTCCAG 610
Qy      148 ProGluSerGlnIleValGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
      611 CCNAATTTCCACTGAGTCCACGAAGTGGAGTGTCTCGATTGTGATGCAAGATTTTCATCC 670
Qy      168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrPro---GluGluPro 186
      671 CACCTCTCTGCAGTCTAACATGGAGTTCATCGACACTCTTACCTATGATGATGAC 730
Qy      187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
      731 AGGATAACTGCCCTACCAACAGGAGTATTGCAGATCTATGATGTTCAGCCAAAGGATTTCT 790
Qy      207 GlySerThrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
      791 GGAAATTTATCGTTGTTGTTGCTGCCACTGTAGCCACCGACGCTAAAGATGTGAGGCGCTCG 850
Qy      227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValIleVal 246
      851 CTAAGTGTGAT-----CCAGCTAAGGAGTCAAAATCCTTCCACACACCAACCAATATA 904
Qy      247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
      905 GCAGGTCCACAGAACATAACAACATCTCTTCATCAGACTGTAGTTTGGATGATGATGCC 964
Qy      267 SerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThr 286
      965 ACAGGAAATCCCAACCAATCATTTCTTGGAGCGCTTGATCACAATATCCATTCATGTC 1024
Qy      287 ---AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
      1025 TTTAATCTCGGTACTTGGAAATGTAATCTCATGATATCTCATGTCAGGCTACACAT 1084
Qy      306 SerGlyValThrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
      1085 GCTGGAGTATATGTTTGTGGGCCACTACCCCTGGCCACACGCAACTTTACAGTTGCTATG 1144
Qy      326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
      1145 GCNACTTTAACTGTATTAGTCTCTCTCTTCAATTTGTTGAATGGCCAGAAAGTTTAAACAG 1204
Qy      346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
      1205 CCTCGAGCTGGCAGCTGCTCGATTTGTGTGAGGAGGAAGTAATCCCTCTCCCAAGATG 1264
Qy      366 HisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGlyGly 385
      1265 TCATGTTGAAAATGGAGGAGGAGATACATTCGAATGTTAGATTTAAATGTAC----- 1318
Qy      386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyThrGlnCysVal 405
      1319 AACAGTAAATGGTAATTAACCCAGATTATCTCGAAGATGATGCTATTATTATCAGTGCATG 1378
Qy      406 AlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgGlu 425
      1379 GCTGAGATAGCCCAAGGATCTATTATATCTAGAGCCAGACTGATGTATGTATGTCAGAA 1438
Qy      426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerSerValLeu 445
      1439 GACAGACCCAGTCTCTCTATATGTCATGCTGAACCATGTCAGCTCAGCTCAGCCATCTT 1498
Qy      446 ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisThr 465
      1499 TTAGCTGGAGAGGCCACTTTTATAATTCAGACAAAGTCATTCCTCTATTCTGTACACTAC 1558
Qy      466 GlnLysAlaArgGlyValAspAsnValGluThrGlnPheAlaValAspAsnAspThrThr 485
      1559 ATGAAAGCAGAGGTTTAAATAATGAAGATATCAAGTAGTATCATCGGAAATGACACAACT 1618

```

Qy	486	GluLeuGlnValAlaArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyr	S05
Db	1619	CATTATATTATGTAGTACTTAGAGCCCTGCCCAAAATTATATTCTTACATTTGTAGCATAT	1678
Qy	506	SerGlnLeuGlyAlaSerA ^g ThrSerSerProAlaLeuValHisThrLeuAspAspVal	525
Db	1679	ATGCCAATGGAGCCCGCAGATGTCTGAACATGTGCACAGAATACTCTAGAGGATGTT	1738
Qy	526	ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleA ^g ValAla	545
Db	1739	CCCTGAGACCTCTCGAAATTAGTTTGACAAGTCGAAGTCCCACATGATTTCTCATCTCC	1798
Qy	546	TripLeuProLeuProSerSerLeuSerSerAsnGlyGlnValLeuTy ^s TyrLysIleGluTyr	565
Db	1799	TGGCTGCCAATCCCGCCAATATCGCGGGGGCAGTGGTGTCTGATCTGCTTGTCTTTC	1858
Qy	566	GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu	585
Db	1859	CGCCTAAGTACTGAGAAATTCATCTCAAGTTCTGGAGCTCCCGGGACCACGCATAGTAC	1918
Qy	586	ThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGly	605
Db	1919	CTTTTGGAGGCCCTCGAAACCTGCACAGTGTCTACTCTGGTTCGGATTACTGCTGCCACCA	1978
Qy	606	AlaGlyTyrGlyValProSerGlnTrpMetGlnHisA ^g ThrProGlyValHisAsnGln	625
Db	1979	GTGGGGCTGGGAGAGTCATCAGTATGCACTTCACATAGCAGCGCCCAAAGCT--ACAAGC	2035
Qy	626	SerHisValProPheAlaProAlaGluLeuLysValArg--AlaLysMetGluSerLeu	644
Db	2036	GTGAAGCCCTTAAGTCTCCA--GAGTTGCATTTGGAGCCTCTGAACCTGTACCAACAT	2092
Qy	645	ValValSerTrpGlnProProHisProThr--GlnIleSerGlyTyrLysLeuTyr	663
Db	2093	TCTGTGAGGTGGCAGCAAGATGTAGAGCACACAGCTCTATTCAGGGCTACAGCTGTAC	2152
Qy	664	TripGlyValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyA ^g ArgGly	683
Db	2153	TACAAGGAGCAAGCGCAGCAGAGAAAT-----GATCAGACTCTCAGCAGCTCCAGGATGCGTG	2179
Qy	684	AspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyrGluLeu	703
Db	2180	-----GGGCCCATTTTCTTGGATACCAAGGACCTTACTCTATATCTCTC	2221
Qy	704	ThrGlnLeuValProGlyA ^g ArgProTyrGluValLysValAlaPheAsnLysHisGlu	723
Db	2222	AGTGGCTTAGACCCCGAAGAAATATATGTGTAGACTCTCTGGCTTACACACACATAGAC	2281
Qy	724	AspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuPro	743
Db	2282	GATGGCTATCAGGCA-----GATCAGACTCTCAGCAGCTCCAGGATGCGTG	2326
Qy	744	IleGlnArgGly-----ProProLeuProProAlaHisValHisAlaGluSer	759
Db	2327	TCTGTTCTGATCGCATGGTCCCTCTCCACCAACCCACCATCTCTATGGAGGCT	2386
Qy	760	AasnSerSerThrSerIleTrpLeuArgTriplyLysProAspPheThrThrValLysIle	779
Db	2387	AACACCTCATCTTCATCTCTCTGCATCTGGAGGGCTGCATTCACCGCTGCACAAATC	2446
Qy	780	ValAsnTyrThrValArgPheGlyProTripGlyLeuArgAsnAlaSerLeuValThrTyr	799
Db	2447	ATTAACATACACCATCCGCTGTAATCTGTTGGCCTGCAGAATGCTTCTTTGGTTCTGTAC	2506
Qy	800	TyrThrSerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrLysTyrGlu	819
Db	2507	CTTCAACATCAGAAACTCATGTCTGGTTCAAGGCTCTAGAACCCMAACACCAATACGAA	2566
Qy	820	PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg	839
Db	2567	TTTGCCCTTGATATACATGTGGATCAGCTTTCAGTCTCTTGGAGCCCTGTAGTCTACCAT	2626
Qy	840	SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr	859

```

Db      2627  TCTACTCTTCCAGAGCACGACAGGCCACCACGAGTTGGAGTAAAGTGACATTAATAGAG 2686
Qy      860  ProSerThrValArgLeuHisTyrCysProProThrGluProAenGlyGluIleValGlu 879
Db      2687  GATGACACTGCCCTGGTGTCTTGGAAACCCCTGATGGCCACAGAAACAGTTGTGACCCGC 2746
Qy      880  TyrLeuIleLeuTyrSerAenAenHisThrGlnProGluHisGlnTrrpThrLeuLeuThr 899
Db      2747  TATACTATCTTATATGATCTAGAGAAAGCCCTGATATGTCAGAGAGAGTGCCAGGCTTACAC 2806
Qy      900  ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhe 919
Db      2807  CGTGAAGGGCAATACCATGGCTTTGCTAGAAAACTTGTGTACGGAATGTGTACATT 2866
Qy      920  PheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
Db      2867  GTCACAGATATCTGCATCCATGAGTGGGAGAGGAGCCCTTTTCAAATTTCTGTGGAGCTG 2926
Qy      940  IleThrIleuGlnGluThrPheSerAspSer-----LeuAspValHisAlaValThrGlyVileIle 959
Db      2927  GCAGTACTTCCAAAGGAAACCTCTGAAATCAAATCAGAGGCCCAAGCGTTTAGATTCTGCT 2986
Qy      950  -----LeuAspValHisAlaValThrGlyVileIle 959
Db      2987  GATGCCAAAGTTATTTCAGGATATTACCATTCTGCACCAAAAATCAATGATCGCATGCT 3046
Qy      960  ValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln 979
Db      3047  GTAGGTGTGGCATAGCCTTGACTGCCTCATCTGTGTCTCTCTCTGTATATACCGA 3106
Qy      980  SerSerHisArgGluAla-----LeuPro 987
Db      3107  AGTAAAGCCAGGAAATCATCTGCTTCCAGACGGCACAGAACTGAACCTCAACAGTTACCT 3166
Qy      988  GlyLeuSerSerSerGlyThrProGlyAsn 997
Db      3167  CGTACCAGTGCCTCCTTAGCTAGTGGAAAT 3196

RESULT 4
US-09-877-730-11
; Sequence 11, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sander, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Enc
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2976
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-11

Alignment Scores:
Pred. No.: 1,79e-102 Length: 2976
Score: 1691.50 Matches: 375
Percent Similarity: 55.79% Conservative: 160
Best Local Similarity: 39.10% Mismatches: 347
Query Watch: 77
DB: 4 Gaps: 15

```



```
QY 791 LeuArgAsnAlaSerLeuValThrTyrThrSerSerGlyGluAspLeuLeuGly 810
Db 2056 CTGCAGATGCTTCTTGGTCTGTACTCTCAACATCAGAACTCACATGTTGTTCAA 2115
QY 811 GlyLeuLysProPheThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAsp 830
Db 2116 GGTCTAGAACCAACCAACCAATTTGCCGTTTCGATTTACATGTGGATCAGCTTTTCC 2175
QY 831 GlyProPheGlySerValValGluArgSerThrLeuProAspArgProSerThrProPro 850
Db 2176 AGTCTTGGAGCCCTGTAGTCTACCAATTTCTCTTCCAGAGCACACAGAGCCCAACA 2235
QY 851 SerAspLeuArgLeuProLeuThrProSerThrValArgLeuHisTrpCysProPro 870
Db 2236 GTTGGAGTAAAGTACATTAATAGAGGATGACACTGCCCTGTTCTTGGAAACCCCT 2295
QY 871 ThrGluProAsnGlyGluLeuValGluTyrLeuLeuLeuLysSerAsnAsnHisThrGln 890
Db 2296 GATGGCCCAAGAACAGTGTGACCGCTATCTATCTATATGATCATCTAGGAGGCTGG 2355
QY 891 ProGluHisGlnTrpThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHis 910
Db 2356 ATTCGAGGAGGTGGAGGTCTTACACCGTGAAGGGGCAATAACCATGGCTTTGCTAGAA 2415
QY 911 GlyLeuGluSerAspThrArgTyrPhePheLeuMetGlyAlaArgThrGluValGlyPro 930
Db 2416 AACTTGGTAGCAGGAATGTATCATTTGTCAAGATATCTGCATCCAATGAGTGGGAGAA 2475
QY 931 GlyProPheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSer--- 949
Db 2476 GGACCCCTTTCAAAATCTGTGGAGTGGCAGTACTTCCAAAGGAACCTCTGAATCAAAAT 2535
QY 950 -----Leu 950
Db 2536 CAGAGGCCCAAGCGTTAGATTTCTGTGATGCCAAAGTTTATTACAGATATTACCATCTG 2595
QY 951 AspValHisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuLeu 970
Db 2596 GACCAAAATCAATGACTGTCATGCTGTAGTGTGGCATGTGCTTGCATCTGACCTGCATCCTC 2655
QY 971 AlaCysMetCysAlaGlyLeuArgGlnSerSerHisArgGluAla----- 985
Db 2656 ATCTGTGTTCTCATCTTGATATACCGAAGTAAAGCCAGGAATCATCTGCTTCCAAAGACG 2715
QY 986 -----LeuProGlyLeuSerSerGlyThrProGlyAsn 997
Db 2716 GCACAGAAATGGAATCAACAGTATACCTCGTACCGAGTGCCTCTTAGCTAGTGGAAAT 2772
```

RESULT 5

```
; Sequence 17, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3219
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-17
```

Alignment Scores:

```
Pred. No.: 2e-102 Length: 3219
Score: 1691.50 Matches: 375
Percent Similarity: 55.79% Conservative: 160
Best Local Similarity: 59.10% Mismatches: 347
Query Match: 25.61% Indels: 77
DB: 4 Gaps: 15
```

US-09-754-997A-2 (1-1252) x US-09-877-730-17 (1-3219)

```
QY 79 GluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlu 98
Db 7 GAAATAAAGCGATCGAGGTTCCTTTTAAACGCGCTCTTTATACATCAGT-----GAG 57
QY 99 GlnGluAspSerAspGluAlaLeuArgIleTrpLysValThrGluGlySerTyr 118
Db 58 GTGGAAGCGCAGGCGAGGAGCAGTCC-----GATGAAGGATTTTAT 99
QY 119 SerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeu 138
Db 100 CAGTCTTGGCAATGACAAATATATGAGCCATTTCTTAGTCAAAAAGCTCATCTTGCCTTA 159
QY 139 AlaThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThr 158
Db 160 TCAACTATTCTGCAATTTGAAGTCCAGGCCAATTTCCACTGAGGTCCACGAGGTGGAGTT 219
QY 159 AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAsp 178
Db 220 GCTCGATTTGCAATGCAAGATTTATCCCAACCTCTCTGAGTCATATCAATGGAGTTCAT 279
QY 179 GlnValThrValPro---GluGluProArgLeuLeuThrLeuProLysTrpLeuGln 197
Db 280 CGGACACTCTTACTATGACTATGACGAGGATTAATCGCCTACCAACAGGATTTGCGAG 339
QY 198 IleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAla 217
Db 340 ATCTATGATGTGAGCAAGGATTTCTGGAATATATCGTTGTTATTTGCTGCCACTGTAGCC 399
QY 218 ArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAla 237
Db 400 CACCGACGTAAAGTATGAGGCGCTCGCTTAATCTGTGATT-----CCAGCTAAGGAGTCA 453
QY 238 ThrArgGlyGlnAspValValIleValAlaAlaProGluAsnThrThrValValSerGly 257
Db 454 AAATCTCTCCACACACACCAATATATAGCAGTCCACAGACATCAACAATCTCTTCAT 513
QY 258 GlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpVal 277
Db 514 CAGACTGTAGTTTGGATGTCATGCGCACAGGAATCCCAAAACCAATCATCTTCTTGGAGC 573
QY 278 ArgGlnAspGlyLysProIleSerThr---AspValIleValLeuGlyArgThrAsnLeu 296
Db 574 CGCTTGTATCACAATATCCATTCATGTCTTTAATATCTCGGGTACTTGGAAATGTGTAATCTC 633
QY 297 LeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysPro 316
Db 634 ATGATATCTGATGTCTAGGCTACACATGCTGGAGTATATGTTGTGGGCCACTACCCCT 693
QY 317 LeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle 336
Db 694 GGCACACGCAACTTTACAGTTGCTATGCGCAACTTTAACTGTATTAGTCTCTCTCTCATTT 753
QY 337 SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg 356
Db 754 GTTGAATGGCCAGAAAGTTTAAACAGGCTCGAGCTGGCAGTCTGCTGATTTGTGTGTCAG 813
QY 357 AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgPro 376
Db 814 GCAGAAGGATCCCTCTCCAGATGTCTATGTTGAAATATGGAAGGAGATACATTCG 873
QY 377 AsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeu 396
```

874 AATGGTAGAATTAAATGTAC-----AACAGTAAATGGTAAATTAACACAGATTATTCCT 927
Db
397 GlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAla 416
Qy
928 GAAGATGATGCTATTATTCAGTGCATGGCTGAGATAGCAAGGATCTATTTTACTAGA 987
Db
417 AlaProLeuAlaValValAlaGluGlyLeuProSerAlaProThrArgValThrAla 436
Qy
988 GCCAGACTGACTGTAGTGTGTCAGCAAGACAGACAGCCAGTGTCTCCCTATATATGATGCT 1047
Db
437 ThrProLeuSerSerSerValLeuValAlaTyrGluArgProGluLeuHisSerGlu 456
Qy
1048 GAAACCATGTCAAGCTCAGCCATCTTTTAGCTGGGAGGAGCCACTTTTAAATTCAGAC 1107
Db
457 GlnIleGlyPheSerLeuHisTyrGlnLysAlaAaGlyValAspAsnValGluTyr 476
Qy
1108 AAGTCACTTCCTTCTGACATACATGAAAGCAGAGGTTTAAATTAATGAAGTAT 1167
Db
477 GlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr 496
Qy
1168 CAAGTAGTCTCGGAATGACCAACTCATTTATTTATTTGATGACTTAGAGCTGCCAGC 1227
Db
497 AspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerPro 516
Qy
1228 AATTATCTTCTACATTGTAGCATATATGCCAATGGGAGCCAGCCAGATGCTGACCAT 1287
Db
517 AlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSer 536
Qy
1288 GTGACACAGAACTACTAGAGGATGTTCCCTGAGACCTCTCGAAATTAATTTGACAGT 1347
Db
537 ProAsnProSerAspIleArgValAlaTyrLeuProLeuProSerSerLeuSerAsnGly 556
Qy
1348 CGAAGTCCCACTGATATTCCTATCTCTGCTGCCAATCCAGCCAAATATCGCGGGGC 1407
Db
557 GlnValLeuTyrTyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThr 576
Qy
1408 CAAGTGTGCTGTATCGCTGTGCTTTCCGCTTAAGTACTGAGAAATTCATCAAGTTCTG 1467
Db
577 GluValProGlnAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr 596
Qy
1468 GAGTCCCGGGAGCCAGCAGTACCTTTTGGAGGCTGAAACCTGACAGTGTCTAC 1527
Db
597 ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln 616
Qy
1528 CTGTTTCGATTACTGCTGCCACAGAGTGGGCTGGGAGAGTCACTAGTATGGACTTCA 1587
Db
617 HisArgThrProGlyValHisGlnSerHisValProPheAlaProAlaGluLeuLys 636
Qy
1588 CATAGAGCCCAAGCT--ACAAGCGTGAAGCCCTTAAGTCTCCA---GAGTTGCAAT 1641
Db
637 ValArg--AlaLysMetGluSerLeuValValSerTyrGlnProProHisProThr 655
Qy
1642 TTGGAGCCTCTGACCTGATACCATTTCTGTGAGGTGGCAGCAGATGTAGAGGACACA 1701
Db
656 ---GlnIleSerGlyTyrLysLeuTyrTyrGlyValGlyThrGluGluGluAlaAsp 674
Qy
1702 GCTGCTATTTCAGGCTACAGCTGTACTACAAGAGAGAGGCGCAGCAGAGAAT----- 1755
Db
675 GlyAspArgProGlyGlyArgGlyAspGlnAlaTyrAspValGlyProValArgLeu 694
Qy
1756 -----GGGCCCAATTTTCTTG 1770
Db
695 LysLysValValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluVal 714
Qy
1771 GATACCAAGGACTCTCTACTCTCTAGTGGCTTAGACCCAGAGAAATATCATGTG 1830
Db
715 LysLeuValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThr 734
Qy
1831 AGACTCTGCTTACCAACATAGAGATGCTATCAGCA-----GAT 1875
Db
735 GluLysAlaProThrProAspLeuProIleGlnArgGly-----ProProLeu 750
Qy
1876 CAGACTGTGACGACTCCAGATGCGGTCTGTCTGTGATGATGCTGCTCTCTCCACCA 1935
Db

751 ProProAlaHisValHisAlaGluSerAsnSerThrSerIleTrpLeuArgTrpLys 770
Qy
1936 CCACCCCACTCTCTATGCGAAGCTAAACCTCATCTTCCATCTTCTGCTGACCTGGAG 1995
Db
771 LysProAspPheThrThrValLysIleValAsnTyrThrValArgPheGlyProTrpGly 790
Qy
1996 AGGCTGATCTACCGCTGCACAAATCAATTAACATACACCATCCGCTGTAATCTCTGTGC 2055
Db
791 LeuArgAsnAlaSerLeuValThrTyrThrSerSerGlyGluAspIleLeuIleGly 810
Qy
2056 CTGCAGATGTTCTTTGTTCTGTACCTTCAACATCAGAAATCAGATCATGTGTGTTCAA 2115
Db
811 GlyLeuLysProPheThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAsp 830
Qy
2116 GGTCTAGAACCAACACCAATACGAATTTGCGGTTCGATTACATGTGATCAGCTTCC 2175
Db
831 GlyProPheGlySerValValGluArgSerThrLeuProAspArgProSerThrProPro 850
Qy
2176 AGTCTTGGAGCCCTGTAGTCTACCATTTCTTCCAGAGCCAGCAGGCCACCA 2235
Db
851 SerAspLeuArgLeuSerProLeuThrProSerThrValArgLeuHisTyrCysProPro 870
Qy
2236 GTTGAGGTAAAGTGACATTAATAGAGATGACACTGCCCTGGTTCTTGGAAACCCCT 2295
Db
871 ThrGluProAsnGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGln 890
Qy
2296 GATGCCCCAGAAACAGTTGTGACCGCTATCTACTATCTATATGATCTAGGAAGCCTGG 2355
Db
891 ProGluHisGlnTyrThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHis 910
Qy
2356 ATTGAGGAGAGTGGCAGTCTTACACCGTGAAGGGGCAATAACCATGCTTTGCTAGAA 2415
Db
911 GlyLeuGluSerAspThrArgTyrPhePheLysMetGlyAlaAaThrGluValGlyPro 930
Qy
2416 AACTGTGAGCAGGAATGTGTACATTCAGATATTCGATCCATCAGAGTGGAGAA 2475
Db
931 GlyProPheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSer--- 949
Qy
2476 GGACCTTTTCAAAATCTGTGGAGTGGCAGTACTTCCAAAGGAAACCTCTGAATCAAT 2535
Db
950 -----Leu 950
Qy
2536 CAGAGCCCAAGCGTTAGATTCTGCTGATGCCAAAGTTTATTCAGGATATTCACATCTG 2595
Db
951 AspValHisAlaValThrGlyIleValGlyValCysLeuGlyLeuLeuCysLeuLeu 970
Qy
2596 GACCAAAATCAATGACTGGCATTTGCTGTAGTGTGGCATAGCCTTGACCTGCATCTC 2655
Db
971 AlaCysMetCysAlaGlyLeuArgGlnSerSerHisArgGluAla----- 985
Qy
2656 ATCTGTGTCTCTCATCTTGATATACCAAGTAAAGCCAGGAATCATCTGCTTCCAGAGC 2715
Db
986 -----LeuProGlyLeuSerSerSerGlyThrProGlyAsn 997
Qy
2716 GCACAGATGGAACTCAACAGTTACTCGTACCAAGTGGCTCTCTAGTAGTGAAT 2772
Db

RESULT 6

US-09-877-730-5

Sequence 5, Application US/09877730

Patent No. 6465632

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Scoville, John

APPLICANT: Turner, C. Alexander Jr.

APPLICANT: Friedrich, Glenn

APPLICANT: Abulin, Alejandro

APPLICANT: Zambrowski, Brian

APPLICANT: Sands, Arthur T.

TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding t

FILE REFERENCE: LEX-0189-USA

CURRENT APPLICATION NUMBER: US/09/877,730

CURRENT FILING DATE: 2001-06-08


```
Db 1554 ----- 1554
Qy 646 ValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrpGly 665
Db 1554 ----- 1554
Qy 666 GluValGlyThrGluGluGluAlaAspGlyAspArgProProGlyGlyArgGlyAspGln 685
Db 1554 ----- 1554
Qy 686 AlaTrpAspValGlyProValArgLeuLysLysLysValLysGlnTyrGluLeuThrGln 705
Db 1554 ----- 1554
Qy 706 LeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGluAspGly 725
Db 1555 ----- 1602
Qy 726 TyrAlaAlaValTrpLysGlyThrGluLysAlaProThrProAspLeuProIleGln 745
Db 1603 TATCAGGCA-----GATCAGACTGTGCAGACTCCAGGATGGTGTCTGTT 1647
Qy 746 ArgGly-----ProProLeuProAlaHisValHisAlaGluSerAsnSer 761
Db 1648 CGTGATCCAGTCCCTCCACCCACCCACCATCTCTATGCGAGGCTTAACACC 1707
Qy 762 SerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsn 781
Db 1708 TCATCTTCATCTCTCTGATCGAGGAGGCTGCATTCACCGTCGCAATCATTAAC 1767
Qy 782 TyrThrValArgPheGlyProTrpGlyLeuArgAlaSerLeuValThrTyrTrpThr 801
Db 1768 TACACCATCCGCTGTAATCCTGTGGCTGCGAATGCTCTCTGTGTTCTGTACCTCAA 1827
Qy 802 SerSerGlyGluAspIleLeuIleGlyLysLysProPheThrThrLysTrpGluPheAla 821
Db 1828 ACATCAGAACTCATGTTGGTCAAGGCTTAGAACCAACCAACCAATACGAAATTTGCC 1887
Qy 822 ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr 841
Db 1888 GTTCGATTACATGTGATCAGCTTCCAGTCTCTGGAGCCCTGTAGTCTACCATCTACT 1947
Qy 842 LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer 861
Db 1948 CTTCCAGAAGCACCGAGCCACACAGTTCGAGTAAAGTGACATTAATAGAGGATGAC 2007
Qy 862 ThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyrLeu 881
Db 2008 ACTGCCCTGGTTCTTGAAACCCCTCATGCGCCAGAAACAGTGTGACCCGCTATACT 2067
Qy 882 IleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrGlu 901
Db 2068 ATCTTATATGATCATAGAGAGCCCTGGATTGCGAGAGTGGCGAGTCTTACCGTGAA 2127
Qy 902 GlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLys 921
Db 2128 GGGCAATAACCATGGCTTTCTAGAAACCTTGGTAGCAGGAAATGTCTACATGTCAAG 2187
Qy 922 MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
Db 2188 ATATCTGCATCAATGAGTGGAGAGGAGCCCTTTTCAAAATCTCTGTGGAGCTGCGAGTA 2247
Qy 942 LeuGlnGluThrPheSerAspSer----- 949
Db 2248 CTTCCAAAGAAACCTCTGAAATCAATCAGAGGCCCAAGCGTTTATGATTCGTGATGCC 2307
Qy 950 ----- 961
Db 2308 AAGTTTATTTCAGGATATTACATCTGACCAAAATCAATGACTGGCATTTGCTGTAGGT 2367
Qy 962 ValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGlnSerSer 981
Db 2368 GTTGGCATAGGCTTGACCTGCATCTCTGTGTCTCATCTTGTATATACCAAGTAA 2427
```

```
Qy 982 HisArgGluAla-----LeuProGlyLeu 989
Db 2428 GCCAGGAATCATCTGCTTCCAGACGCGCAGCAATGGAATCAACAGTTTACCTCGTACC 2487
Qy 990 SerSerSerGlyThrProGlyAsn 997
Db 2488 AGTCCCTCTCTAGCTAGTGGAAAT 2511
```

RESULT 7

```
US-09-877-730-9
; Sequence 9, Application US/09877730
; Patent NO. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-9
```

```
Alignment Scores:
Pred. No.: 3,56e-82 Length: 2958
Score: 1384.00 Matches: 334
Percent Similarity: 46.11% Conservative: 140
Best Local Similarity: 32.49% Mismatches: 310
Query Match: 20.96% Indels: 244
DB: 4 Gaps: 14
```

US-09-754-997A-2 (1-1252) x US-09-877-730-9 (1-2958)

```
Qy 9 GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGlyLeuProLeuProGln 28
Db 43 GGGATGCTGCTCCGCGCGCTCTGCTCTCTGCTGTCAGT-----CCTTTGCCAGGA 96
Qy 29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValIleLeuGlyPro 47
Db 97 GTGTGGTGTCTTACCGAATGCTCTTTGTAAAGAACCAACAGGATGTAACGTGTACAAGA 156
Qy 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArg 67
Db 157 AAGGACCCAGTCGTTTATTGATTGCCAGCTCAGGAGAGTT-----CCTATTAAAG 207
Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro 87
Db 208 GTCCATGTTGTAATAATGGAGCAAAATGCTGANAATAAACGGATCGAGGTTCTTTCT 267
Qy 88 AsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAspSerAspAspGluGluAla 107
Db 268 AACGGCTCTTTATACATCAGT-----GAGGTGGAGGCGAGGAGGAGCAGTCC 318
Qy 108 LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly 127
Db 319 -----GATGAAGGATTTTATCAGTCTTGGCAATGCAACAAATATGGA 360
Qy 128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
Db 361 GCCATTCTTAGTCAAAAGGCTCATCTTGCCTTATCACTATTCTTCGCAATTTTGAAGTCCAG 420
```



```
Qy 457 GlnIleIleGlyPheSerLeuHisTyrGlnIysAlaArgGlyValAspAsnValGluTyr 476
Db 1108 AAGTCATTGCTTATTCGTACACTATACGAAGCAGAGGTTTAAATAATGAAGAGTAT 1167
Qy 477 GlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr 496
Db 1168 CNAAGTAGTCGGAATATGACACAACTCATATATATTATGATGACTTAGAGCCCTGCCAGC 1227
Qy 497 AspTyrGluPheTyrValAlaTyrSerGlnLeuGlyAlaSerArgThrSerPro 516
Db 1228 AATTATACTTTTACATTGTAGCATATATGCAATGCGAGCCAGACGATGCTGCACCAT 1287
Qy 517 AlaLeuValHisThrLeuAspValProSerAlaAlaProGlnLeuThrLeuSerSer 536
Db 1288 GTGACACAGAACTACTAGAGATGACCCAGCA----- 1320
Qy 537 ProAsnProSerAspIleAlaGValAlaTriLeuProLeuProSerSerLeuSerAsnGly 556
Db 1320 ----- 1320
Qy 557 GlnValLeuLysTyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThr 576
Db 1320 ----- 1320
Qy 577 GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr 596
Db 1320 ----- 1320
Qy 597 ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTriMetGln 616
Db 1320 ----- 1320
Qy 617 HisArgThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLys 636
Db 1320 ----- 1320
Qy 637 ValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGln 656
Db 1320 ----- 1320
Qy 657 IleSerGlyTyrLysLeuTyrTriPtyGluValGlyThrGluGluAlaAspGlyAsp 676
Db 1320 ----- 1320
Qy 677 ArgProProGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLys 696
Db 1320 ----- 1320
Qy 697 LysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeu 716
Db 1321 -----AGAAAATATCATGTGAGACTC 1341
Qy 717 ValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTriPtyAsGlyLysThrGluLys 736
Db 1342 CTGGCTTACAACAACATAGACATGGCTATCAGGCA-----GATCAGACT 1386
Qy 737 AlaProProAspLeuProIleGlnArgGly-----ProProLeuProPro 752
Db 1387 GTCAGCACTCCAGGATGCTGTCTGTCGTGATGCATGGTCCCTCCACCACCAACC 1446
Qy 753 AlaHisValHisAlaGluSerAsnSerSerThrSerIleTriPtyLeuArgTriPtyLysPro 772
Db 1447 CACCATCTCTATGCGAAGGCTAAACACCTCATCTTCATCTTCCTGCACGTGAGGAGGCT 1506
Qy 773 AspPheThrValLysIleValAsnTyrThrValArgPheGlyProTriPtyGlyLeuArg 792
Db 1507 GCATTACCGCTGCACAATCATTTAATACCATCCGCTGTAACTCTGTGGCTGAG 1566
Qy 793 AsnAlaSerLeuValThrTyrThrSerSerGlyGluAspIleLeuIleGlyGlyLeu 812
Db 1567 AATGCTCTTTGGTTCTGTACCTTCAACATCAGAAACCTCACATGTTGTTCAAGGTCTA 1626
Qy 813 LysProPheThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyPro 832
```

```
Db 1627 GAACCAAAACCAACAAATACGAATTTGCGTTCGATTACATGTGGATCAGCTTTCCAGTCT 1686
Qy 833 PheGlySerValValGluArgSerThrLeuProAspArgProSerThrProProSerAsp 852
Db 1687 TGGAGCCCTGTAGTCTACCATTTCTACTTCTCCAGAAGCACACGAGGCCCCACCAAGTTGGA 1746
Qy 853 LeuArgLeuSerProLeuThrProSerThrValArgLeuHisTyrCyAspProThrGlu 872
Db 1747 GTAAAGTAGTACATTAATAGAGATGACACTGCCCTGGTTTCTTGGAAACCCCTGATGCG 1806
Qy 873 ProAsnGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGlu 892
Db 1807 CCAGAAACAGTTGTGACCGCTATATCTTATATGTCATCTAGGAAGSCCTGGATTGCA 1866
Qy 893 HisGlnTriPtyThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeu 912
Db 1867 GGAAGTAGTGGCAGTCTTACACCGTGAAGGGGCAATAACCATGCTTGTCTAGAAAACTTG 1926
Qy 913 GluSerAspThrArgTyrPhePheLysMetGlyAlaArgThrGluValGlyProGlyPro 932
Db 1927 GTAGCAGGAAATGTGTACATTTGTCAAGATATCTGCATCCATGAGGTGGAGAGGCC 1986
Qy 933 PheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSer----- 949
Db 1987 TTTTCAATCTGTGGAGCTGGCAGTACTTCCAAAGGAAACCTCTGAAATCAATCAGAGS 2046
Qy 950 -----LeuAspVal 952
Db 2047 CCCAAGCGTTAGATTCTGCTGATGCCAAAGTTTATTTCAGGATATTACCATCTGGACCA 2106
Qy 953 HisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuAlaCys 972
Db 2107 AAATCAATGACTGGCATGCTGTAGGTGTGGCATAGCTTGACCTGCATCTCATCTGT 2166
Qy 973 MetCysAlaGlyLeuArgGlnSerSerHisArgGluAla----- 985
Db 2167 GTTCTCATCTTGATATACCGAAGTAAAGCCAGGAATCATCTGCTTCCAAGCGGCACAG 2226
Qy 986 -----LeuProGlyLeuSerSerSerGlyThrProGlyAsn 997
Db 2227 AATGGAACTCAACAGATTACCTCGTACAGTGCCTCTTAGTAGTGGAAAT 2277
```

RESULT 9

```
US-09-877-730-19
; Sequence 19, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 2724
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-19
```

```
Alignment Scores:
Pred. No.: 1.36e-78 Length: 2724
Score: 1329.00 Matches: 312
Percent Similarity: 46.29% Conservative: 131
```

Best Local Similarity: 32.60%		Matches: 276	
Query Match: 20.12%		Indels: 238	
DB: 4		Gaps: 11	
US-09-754-997A-2 (1-1252) x US-09-877-730-19 (1-2724)			
Qy	79	GluHisGluAenLeuHisLeuLeuProAenGlySerLeuTrpLeuSerSerProLeuGlu	98
Db	7	GAATAAAGCGATCAGGCTCTTTCTAAACGGCTCTTTATACATCAGT-----GAG	57
Qy	99	GlnGluAenSerAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTyr	118
Db	58	GTGAAGCGCGAGGAGCAGTCC-----GATGAAGGATTTTAT	99
Qy	119	SerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeu	138
Db	100	CAGTGTGGCAATGAACAAATATGGAGCCATCTTAGTCAAAAGCTCATCTTGCCTTA	159
Qy	139	AlaThrLeuGluAenPheSerLeuHisProGluSerGlnIleValGluGluAenGlyThr	158
Db	160	TCAACTATTTCTGCATTTGAAGTCCAGCCCAATTTCCACTGAGGTCACGAAGGTGAGTT	219
Qy	159	AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAsp	178
Db	220	GCTCGATTTCATGCAAGATTTCATCCACCCTCTCTGCAGTCATAACATGGAGTTCAAT	279
Qy	179	GlnValThrValPro---GluGluProArgLeuIleThrLeuProLysTrpLeuGln	197
Db	280	CGGCAACTCTACCTATGACTATGGACAGATAACTGCCCTTACCACAGGAGTATGGAG	339
Qy	198	IleLeuAenValGlnAenSerAspAlaGlySerTyrArgCysValAlaAlaThrAsnSerAla	217
Db	340	ATCTATGATGTCAGCCAAAGGATTCGGAATTTATCGTTATGCTGCCACTGTAGCC	399
Qy	218	ArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAla	237
Db	400	CACCGACGTAAGATGATGAGCGCTCGCTAACTGTGATT-----CCAGCTAAGGAGTCA	453
Qy	238	ThrArgGlyGlnAenValValIleValAlaAlaProGluAenThrThrValValSerGly	257
Db	454	AAATCCTTCCACACACCAACCAATTTATAGCAGGTCCACAGAACATAACCAATCTTTCAT	513
Qy	258	GlnAenValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpVal	277
Db	514	CAGACTGTAGTTTGGATGCTATGGCCACAGGAATCCCAACCAATCATTTCTTGGAGC	573
Qy	278	ArgGlnAenGlyLysProIleSerThr---AspValIleValLeuGlyArgThrAsnLeu	296
Db	574	CGCCTTGATCACAATTCATTCATGCTTTTAATACTCGGGTACTTGGAAATGGTAATCTC	633
Qy	297	LeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAenLysPro	316
Db	634	ATGATATCTGATGTCAGGTACCAACATGCTGGAGTATATGTTGTCGGGCCACTACCCCT	693
Qy	317	LeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle	336
Db	694	GGCACAGCACTTTACAGTTGCTATGGCACTTTAACTGTATTAGTCTCTCTTCATTT	753
Qy	337	SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg	356
Db	754	GTTGAATGGCCAGAAAGTTTAAACAGGCCCTCGAGCTGCGACTGCTCGATTGTGTGTCAG	813
Qy	357	AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgPro	376
Db	814	GCAGAAGGAATCCCTCTCCCAAGATGTCATGGTTGAAAAATGGAAGGAACATACATTCG	873
Qy	377	AsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeu	396
Db	874	AATGGTAGAATTAAGATGTAG-----AACAGTAAATTTGGTAATTAACAGATTATTCTCT	927
Qy	397	GlnAspAlaGlyTyrTyrGlnCysValAlaGluAenSerAlaGlyThrAlaCysAlaAla	416
Db	928	GAAGATGATGCTATTTATCAGTGTGATGCTGAGATAGCCAGGATCTATTATTTATCTAGA	987

773 AspPheThrValValLysIleValAsnThrValArgPheGlyProTrpGlyLeuArg 792
1507 GCATTCACCGCTGCACAAATCATTAACATCACCATCGCTGTAACTCTGTGGCTTCGAC 1566
793 AsnAlaSerLeuValThrTyrThrSerSerGlyGluaspIleLeuLeuGlyLeu 812
1567 AATGCTCTCTTGTGTCTGTACCTTCAACATCAGAACTCACATGTGTGTTCAGGTCTA 1626
813 LysProPheThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyPro 832
1627 GAACCAACACCAAAATACGAATTTGCCGTTTCGATTACATGTCATCAGCTTTCCACGCTCT 1686
833 PheGlySerValValGluArgSerThrLeuProAspArgProSerThrProProSerAsp 852
1687 TGGAGCGCTCTAGCTACCATTTCTACTCTTCCAGAAGCACACAGAGCCGCCACCATGTGA 1746
853 LeuArgLeuSerProLeuThrProSerThrValArgLeuHisTyrCysProProThrGlu 872
1747 GTAAAGTAGCATTAATAGAGGATGACATGCGCTGTCTTCTTGGAAACCCCTGATGCG 1806
873 ProAsnGlyLeuIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGlu 892
1807 CCAGAACACAGTTGTGACCGCTATATCTTATATGTCATCTAGGAAGGCTGGATTGCA 1866
893 HisGlnTrpThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeu 912
1867 GAGAGTGGCAGGTCTTACACCGTGAAGGGCAATAACCATGGCTTTGCTAGAAAACTTG 1926
913 GluSerAspThrArgTyrPhePheLysMetGlyAlaArgThrGluValGlyProGlyPro 932
1927 GTAGCAGGAATGTGTACATTTGTCAAGATATCTGCATCCAAATGAGTGGGGAAGGACCC 1986
933 PheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSer----- 949
1987 TTTTCAAAATCTGTGGAGCTGGCAGTACTTCCAAAGGAACCTCTGAATCAATCAGAGG 2046
950 -----LeuAspVal 952
2047 CCCAAGCGTTTGTAGATCTGCTGATGTCGCAAGTTTATTTCAGGATATTACCATCTGGACCA 2106
953 HisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuValAlaCys 972
2107 AAATCAATGACTGTCATGTCTGTAGGTGTGGCATAGCCTTGACCTGCATCTCATCTGT 2166
973 MetCysAlaGlyLeuArgGlnSerHisArgGluAla----- 985
2167 GTTCTCATCTTGATATACCGAGTAAAGCCAGGAAATCATCTGTTCCAGAGCGGCACAG 2226
986 -----LeuProGlyLeuSerSerSerGlyThrProGlyAsn 997
2227 AATGGAACCTCAACAGATTACTCGTACAGTGCCTCTTAGCTAGTGGAAAT 2277

RESULT 10
US-09-877-730-21
; Sequence 21, Application US/0987730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 21

829 ATTCTGTGGTGGCAGCAGATGTAGAGACACAGCTGCTATTTCAGGGCTACAGCTG 888
Qy TyrTrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProGlyGlyArg 682
Db TACTACAGGAAGAGGAGCAGCAGAAAT----- 918
Qy GlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysValLysGlnTrpGlu 702
Db -----GGGCCCATTTTCTTGGATACCAAGGACCTACTCTATACT 957
Qy LeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHis 722
Db -----CTCAGTGGCTTAGACCCAGAGAATAATCATGTGAGACTCTGGCTTACACACATA 1017
Qy GluAspGlyTyrAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeu 742
Db -----GATCAGACTGTGAGCAGCTCCAGGATGC 1062
Qy ProLeuGlnArgGly-----ProProLeuProAlaHisValHisAlaGlu 758
Db -----GGTCTGTGGTGGATGGCGTCTCTCCACCAACCCACCATCTCTATCGAAG 1122
Qy SerAsnSerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLys 778
Db -----GCTAACACTCATCTTCCATCTTCTGCTGAGGAGGCGCTGCAATTCACGCTGCACA 1182
Qy IleValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThr 798
Db -----ATCATTAATACTACACCATCGCTGTAATCTCTGTGGCTGCAGATGCTTCTTTGGTCTG 1242
Qy TyrTyrThrSerSerGlyGluAspIleLeuLysLysLeuLysProPheThrThrLysTyr 818
Db -----TACCTTCAACACATCAAGAACTCACATGTTGGTTCAAGGTCTAGAACCAACCAATATAC 1302
Qy GluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGlu 838
Db -----GAAATTTGGTGGTTCGATTATGATGATGATGATGATGATGATGATGATGATGATGAT 1362
Qy ArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeu 858
Db -----CATCTACTTCTCCAGAACACACAGCAGCCAGCCAGCTGGAGTAAAGTGCATTAATA 1422
Qy ThrProSerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluLeuVal 878
Db -----GAGGATGACATCGCCCTGCTTTCTGGAAACCCCTGATGGCCAGAACAGTGTGACC 1482
Qy GluTyrLeuIleLeuTyrSerAsnHisThrGlnProGluHisGlnTrpThrLeuLeu 898
Db -----CGCTATATCTTATATGCACTAGGAGGCGCTGGATTCGAGGAGATGGCAGGTCTTA 1542
Qy ThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyr 918
Db -----CACCGTGAAGGGCAATAACCATGCTTGTCTAGAAAACCTGGTAGCAGGAATGTGTAC 1602
Qy PhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAsp 938
Db -----ATTGTCAAGATCTGCAATCAATCAGGTGGGAGAGGAGCCCTTTCAAAATTCGTGGAG 1662
Qy ValIleThrLeuGlnGluThrPheSerAspSer----- 949
Db -----CTGGCAGTACTTCCAAAGGAACCTCTGAATCAAAATCAGAGCCCAAGCGTTTAGATTCT 1722
Qy -----LeuAspValHisAlaValThrGlyIle 958
Db -----GCTGATGCCAAGTTTATTCAGGATATTACCATCTGGACCAAAATCAATGACTGGCAAT 1782
Qy IleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArg 978
Db -----GCTGTAGGTGGTGGATGACCTTGACCTGCATCCTCATCTGTGTTCTCATCTTGTATAC 1842
Qy GlnSerSerHisArgGluAla-----Leu 986
Db -----CGAAGTAAAGCAGGAAATCATCTGCTTCCAGACGGCACAGAAATGGAACCTCAACAGTTA 1902

Qy 987 ProGlyLeuSerSerSerGlyThrProGlyAsn 997
Db 1903 CCTGTACCAAGTGCCTCTTCTAGCTAGTGAAT 1935

RESULT 11

US-09-877-730-27
; Sequence 27, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 2382
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-27
Alignment Scores:
Pred. NO.: 1.85e-66 Length: 2382
Score: 1144.00 Matches: 252
Percent Similarity: 54.69% Conservative: 115
Best Local Similarity: 37.56% Mismatches: 240
Query Match: 17.32% Indels: 64
Gaps: 10

US-09-754-997A-2 (1-1252) x US-09-877-730-27 (1-2382)

Qy 365 LeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGly 384
Db 1 ATGTCATGTTTGAATAATGGAAGGAGATACATTCGAATGGTAGAATTTAAATATGAC--- 57
Qy 385 GlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCys 404
Db 58 ---AACAGTAATTTGGTAATTAACCAAGATTATCTCTGAAGATGATGCTATTATTATCAGTGC 114
Qy 405 ValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArg 424
Db 115 ATGGCTGAGATAGCCAGGATCTATTATTATCTAGAGCCAGACTGACTGTAGTGTATGCA 174
Qy 425 GluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerVal 444
Db 175 GAAACAGACACCAAGTGTCTCCCTATATATGTATACATGCTGAACCACTGTCAAGCTCAGCCATT 234
Qy 445 LeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHis 464
Db 235 CTTTATGCTGGAGAGGCCACTTTTAAATTCAGCAAGTCAATTCCTATTCTGTATCAGC 294
Qy 465 TyrGlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAspThr 484
Db 295 TACATGAAGCAGAGGTTTAAATTAATGAAGATATCAAGTAGTCACTCGGAATGACACA 354
Qy 485 ThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAla 504
Db 355 ACTCATTAATATTATGATGACTTAGAGCCTCCAGCAATATTACTTTCTACATGTAGCA 414
Qy 505 TyrSerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHisThrLeuAspAsp 524
Db 415 TATATGCCAATGGGAGCCAGCCAGATGCTGACCATGTGACACAGAACTACTCTAGAGAT 474

LOCATION: 1...4342
US-09-041-886-24

Alignment Scores:

Pred. No.: 8,57e-53
Score: 943.00
Percent Similarity: 36.29%
Best Local Similarity: 24.54%
Query Match: 14.28%
DB: 3

Length: 4608
Matches: 357
Conservative: 171
Mismatch: 547
Indels: 381
Gaps: 47

US-09-754-997A-2 (1-1252) x US-09-041-886-24 (1-4608)

QY 51 valValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArgValThrTrp 70
DB 169 GTCTCTCTCGACTGCTCC---GGGAGTCCGACCGAGGAGTTCCAGTG---ATCAAGTGG 222
QY 71 SerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeu-----85
DB 223 AAGAAAGATGGC-----ATTCACTGGCCTTGGGAATGGAT 258
QY 86 -----LeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGluGlu 100
DB 259 GAAAGGAGGACGACCACTTTCAATGGTCTCTGCTGATACAAACATCTTCAATCC---315
QY 101 AspSerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTrpCys 120
DB 316 -----AGACACCAAGCCAGATGGGAGCTTTTACCAATGT 351
QY 121 LeuAlaHis---SerProLeuGlyValValAlaSerGlnValAlaValLysLeuAla 139
DB 352 GAGGCATCTTTAGGAGATCTGGCTCAATATTATTAGTGGACACAAAGTTGCGATGCA 411
QY 140 ThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAla 159
DB 412 GGACCACTGAGTTCCTTTTACAGACAGATCTGTACAGCCTTCATGGAGACAGATG 471
QY 160 ArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGln 179
DB 472 CTACTCAAGTGTGAAGTCAATTTGGGAGCCCATGCCAACATCCACTGGCAGGAACCAA 531
QY 180 -----ValThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeu 196
DB 532 CAAGACCTGACTCCAATCCAGATGACTCCGAGTGGTGGTCTTGGCCCTCTGGAGCATTTG 591
QY 197 GlnIleLeuAspValGlnAspSerAlaGlySerTrpArgCysValAlaThrAsnSer 216
DB 592 CAGATCAGCGGCTCCACACGGGGGACATTGGAAATTTACGATGCTCAGCTCGAATCCA 651
QY 217 AlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGlu 236
DB 652 GCCAGCTCAAGACAGCAAGTGAAGCAGAGATTTTATCAGATCCAGGACTGCAT 711
QY 237 AlaThrArgGlyGlnAspValValIleValAlaProGluAsnThrThrValValSer 256
DB 712 -----AGACAGCTGTATTTTCTGCAAGACCATCCCAATAGTAGCCATTGAA 759
QY 257 GlyGlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrp 276
DB 760 GGAAGAGATGCTGCTCGATATGTTGTTCTTGGCTATCTCCCAAGATTTTACCTGG 819
QY 277 ValArgGlnAspGlyLysProIle-----SerThrAspValIleValLeuGly 292
DB 820 TTACGA-----GGCGAGGAGTCATCCAACTCAGGTCTAAAAAGTATTCTTTATTTGGGT 873
QY 293 ArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHisSerGlyValTrpValCysArg 312
DB 874 GGAAGCAACTTGTCTATCTCAATGTGACAGATGATGACAGTGGAAATGTATACCTGTGT 933
QY 313 AlaAsnLysProLeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAla 332
DB 934 GTCACA-----TATAAAATGAGATATTAGTGGCTCTGACAGACTCAGTCTTTGGTT 987

QY 333 AlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArg 352
DB 988 CGCCCATGGTGTAAATCATCTTCCAACTGTATGCTATGCAAGCATGGATATTGAG 1047
QY 353 PheValCysArgAlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIle 372
DB 1048 TTGTAATGTACAGTCTCTGGAAGGCTGTGCCACCTGTGAATGGATGAAGATGAGAT 1107
QY 373 ProLeuArgProAsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThr 392
DB 1108 GTGGTCATTCCTAGTATTATTTTCAGATAGTG---GGAGGAAGCAACTTACGGATCTT 1164
QY 393 GlnIleGlyLeuGlnAspAlaGlyTrpTrpGlnCysValAlaGluAsnSerAlaGlyThr 412
DB 1165 GGGGTGGTGAAGTCAAGTCAAGTCAAGGCTTTTATCAATGTGTGGCTGAAATGAGGCTGAAAT 1224
QY 413 AlaCysAlaAlaAlaProLeu-----AlaValValValArgGluGlyLeu 427
DB 1225 GCCCAGACCGAGTGCACAGTCAATGTCCTAAGCTGCAATCCCAAGCTCCAGTGTCTCT 1284
QY 428 ProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeuValAla 447
DB 1285 CCTTCGGCTCCAGAGATGTGGTCCCTGTCTTGGTTTCCAGCGGATTTGTCCGCTCAGC 1344
QY 448 TrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTrpGlnLys 467
DB 1345 TGGCCGCCACCTCGAGAAGCGAAGGAAACATTTCAAACTTTTCCAGTCTTTTTC 1398
QY 468 AlaAlaGlyValAspAsnValGluTrpGlnPheAlaValAsnAsn-----AspThr 484
DB 1399 TCCAGAGAGGTGACACACAGGGAACGA-----GCATTGAATACACACAGCGCTGGGTCC 1452
QY 485 ThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPheTrpValValAla 504
DB 1453 CTTAGCTCAGTGTGGAAACCTGAAGCCAGAGCCATGTACACCTTTTCAGTGTGGGT 1512
QY 505 TyrSerGlnLeuGlyAlaSerArgThrSerPro-----516
DB 1513 TACAATGAATGGGACCGGGAGAGAGTCTTCAACCCATCAAGGTGGCCACACAGCGCTGAG 1572
QY 516 -----516
DB 1573 TTGCAAGTTCCAGGGCCAGTAGAANAACCTGCAAGCTGTATCTACCTACCTACCTCAATT 1632
QY 516 -----516
DB 1633 CTTATTACCTGGGAACCCCTGCTTATGCAACCGTCCAGTCCAGGTTTACAGATTGTTTC 1692
QY 516 -----516
DB 1693 TGCACCTGAGGTGTCCACAGAAAAGAACAGATATAGAGTTGTAGGACTATCTTATAAA 1752
QY 516 -----516
DB 1753 CTGGAAGGCTGAAAAAATTCACCAATATAGTCTTCCGATTTCTTAGCTTATAATCGCTAT 1812
QY 517 -----AlaLeuValHisThrLeuAspValProSerAla 528
DB 1813 GGTCCGGGGTCTCTACTGATGATATAACAGTGGTTTACACTTCTGAGTGCAGAGTGC 1872
QY 529 AlaProGln---LeuThrLeuSerSerProAsnProSerAspIleArgValAlaTrpLeu 547
DB 1873 CGGCTCAGAACGTCTCCCTGGAGTGGTCAATTCAGAGAGTATCAAGTTAGTGGCTG 1932
QY 548 ProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTrpLysIleGluTrpGlyLeu 567
DB 1933 CCTCTCCATCAGGAACACAAAATGGATTTTATTACCGGCTATAAAATTCGACACAGAAAG 1992
QY 568 GlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeuThrLeu 587
DB 1993 ACGACCCGCGGGGTGAGATGGAAACACTGGAGCCAAACCACTCTGGTACCTA---TTC 2049
QY 588 AsnSerLeuGlnProAsnLysValTrpArgValArgIleSerAlaGlyThrGlyAlaGly 607

Qy 1215 oSerThrAlaIleGluAla-----1222
 Db 3996 CACAGCTGTGTCGACCAACTCACCCACTCCGCGAGCTTTGCTAATCCTTTGCTACCTCC 4055
 Qy 1223 -ProGlySerCysGlnProLysAlaLeuCysProLeuThr-----Va 1237
 Db 4056 ACCAATGAGTCAATAGAACCGAAGTC-----CCTTACACACCACTTTTGTCTCAGCC 4109
 Qy 1237 lSerProSerProArgAlaProValSerSerAlaGlnVal 1251
 Db 4110 AGGCCCACTCTTCTTAAGACCCATGTGAAACAGAGCTCCCTT 4152

RESULT 13

PCT-US94-05277-1
 ; Sequence 1, Application PC/TUS9405277
 ; GENERAL INFORMATION:
 ; APPLICANT: Bruskin, Arthur
 ; APPLICANT: Jarosz, David E.
 ; APPLICANT: Johnson, Karen
 ; APPLICANT: Kinzler, Kenneth W.
 ; APPLICANT: Vogelstein, Bert
 ; APPLICANT: Zabrecky, James R.
 ; TITLE OF INVENTION: Antibodies Specific for DCC Gene Product
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Banner, Birch, McKie & Beckett
 ; STREET: 1001 G Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20001

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/05277
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kagan, Sarah A.
 ; REGISTRATION NUMBER: 32,141
 ; REFERENCE/DOCKET NUMBER: 01107.42709
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202.508.9100
 ; TELEFAX: 202.508.9299
 ; TELEX: 197430 BMB UT
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4608 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORGANISM: Homo sapiens
 ; POSITION IN GENOME:
 ; CHROMOSOME/SEGMENT: 18q21
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..4342
 ; PCT-US94-05277-1

Alignment Scores:
 Pred. No.: 8.57e-53
 Score: 943.00
 Percent Similarity: 36.29%
 Best Local Similarity: 24.54%
 Query Match: 14.28%
 Length: 4608
 Matches: 357
 Conservative: 171
 Mismatches: 547
 Indels: 381

DB: 5 Gaps: 47
 US-09-754-997A-2 (1-1252) x PCT-US94-05277-1 (1-4608)
 Qy 51 ValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArgValThrTyr 70
 Db 169 GTCTCTCTCCAGCTGCTCC---GCGAGTCCGACGAGGAGTTCAGTG---ATCAAGTGG 222
 Qy 71 SerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeu-----85
 Db 223 AAGAAAGATGCG-----ATTCACTGTGGCTTGGGAATGGAT 258
 Qy 86 -----LeuProAsnGlySerLeuTyrLeuSerSerProLeuGluGlnGlu 100
 Db 259 GAAAGGAAGCAGCAACTTTCAATGGTCTCTGTCATACAAACATACATCTTCATTC--- 315
 Qy 101 AspSerAspAspGluGluAlaLeuArgIleTyrLysValThrGluGlySerTyrSerCys 120
 Db 316 -----AGACACCAACAGCCAGATGAGGAGCTTTACCAATGT 351
 Qy 121 LeuAlaHis---SerProLeuGlyValValAlaSerGlnValAlaValValLeuAla 139
 Db 352 GAGGCATCTTTAGGAGATTCGCTCAATTATTAGTCGGACAGCAAAAGTTGCAGTAGCA 411
 Qy 140 ThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAla 159
 Db 412 GGACCACTGAGGTTCTCTTCCACAGACAGAACTCTGCACAGCTTCATGGGAGACACAGTG 471
 Qy 160 ArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTyrGluLysAspGln 179
 Db 472 CTACTCAAGTGTGAAGTCATTGGGAGCCCATCCCAACATCCATGCGAGAAACCAA 531
 Qy 180 -----ValThrValProGluProArgLeuIleThrLeuProLysTyrLeuLeu 196
 Db 532 CAAGACCTGACTCCATCCCGAGTACTCCGAGTGGTGTCTTGGCTCTCGAGAGCATGT 591
 Qy 197 GlnIleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSer 216
 Db 592 CAGATCAGCCGACTCCCAACGGGGGACATTGGAAATTTACCGATGCTCAGCTCGAAATCCA 651
 Qy 217 AlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGlu 236
 Db 652 GCCAGCTCAAGAACAGGAAATGAAGCAGAAATGAGAAATTTTATCAGATCCAGGACTGCAT 711
 Qy 237 AlaThrArgGlyGlnAspValValIleValAlaProGluAsnThrThrValValSer 256
 Db 712 -----AGACAGCTGTATTTCTCAAGACCATCCCAATGTAGTACCATTTGAA 759
 Qy 257 GlyGlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTyr 276
 Db 760 GGAAGAGATGCTGCTCGGAATGTTGTGTTCTGGCTATCTCCACCAAGTTTACCTGG 819
 Qy 277 ValArgGlnAspGlyLysProIle-----SerThrAspValIleValLeuGly 292
 Db 820 TTACGA-----GCGCAGGAGATCATCCAACTCAGGTCTAAAAAGTATCTTTATTGGGT 873
 Qy 293 ArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArg 312
 Db 874 GGAAGCAACTTGCTTATCTCCAATGTGACAGATGATGACATGGAATGTATCTGTGTT 933
 Qy 313 AlaAsnLysProLeuThrArgAspPheAlaThrAlaAlaAlaGluLeuArgValLeuAla 332
 Db 934 GTCACA-----TATAAAATGAGATATTAGTGCTCTCGAGAGCTCACAGTCTTGTT 987
 Qy 333 AlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArg 352
 Db 988 CCGCATGGTTTTTAAATCATCTTCCAACTGTATGCTTATGAAGCATGATATTGAG 1047
 Qy 353 PheValCysArgAlaSerGlyGluProArgProAlaLeuHisTyrLeuHisAspGlyIle 372
 Db 1048 TTTGAATGTACAGTCTCTGGAAGCCCTGTGCCACTGTGAATGGATGAAGAATGGAGAT 1107
 Qy 373 ProLeuArgProAsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThr 392


```
Qy 939 VallleThrLeuGlnGlu-----Thr 945
Db 3214 AGAAGCACCCCTAAATGAGCGGCCAATTGGACAAATGACACCCCGCCATGGCAGTGTCACT 3273
Qy 946 PheSerAspSerLeuValHisAlaValThrGlyValIleValGlyValCysLeuGly 965
Db 3274 CQTCAGAAAGACAGCAACCTGCTGTGATCATTTGTGTGTCACCGTGTGTCTACACAGTG 3333
Qy 966 LeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSerSerHisArgGluAla 985
Db 3334 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3393
Qy 986 LeuProGlyLeuSerSerGlyThrProGlyAsnProAlaLeuThrThrArgAlaArg 1005
Db 3394 AAACGGGCCACCCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3441
Qy 1006 LeuGlyProProSerValProAlaAlaHis---GluLeuGluSerLeuValHisProArg 1024
Db 3442 CTCCGACCCCTGATCTTTGTGATCCTCATGAAAGAAATGGAGATCGAAATAATATTGAAAG 3501
Qy 1025 ProGlnAspTyrSerProProProSerAspValGluAspLysAlaGluValHisSerLeu 1044
Db 3502 CCATCTGGCAGTACCTGCGAGGAGGAC----- 3531
Qy 1045 MetGlyGlySerValSerAspCysArg-----GlyHisSerLys-ArgLys 1059
Db 3532 -----TCTCCATCCAAGTTGCCACAGCCTCACACCAGTCAGCCACAGCAGTCAGAA 3585
Qy 1059 sileSerTrpAlaGln-----Al 1065
Db 3586 ACCCACTGGGAAGCAAAAGCACCTCTCATTCAGTCAAGACACTGAGGAAGCAGGAGC 3645
Qy 1065 aGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuProGln----- 1081
Db 3646 TCTATGTCCACTCTCGAGAGGTGCTGCTGTCGACGCCGAGCC-CCCGCGGCCAAGCTCAT 3704
Qy 1082 -----GlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProAla 1097
Db 3705 GATTCCTCATGAGTCCAGTCCACATCTCTGCTGTGTGAGCGCATCCCGTGCCAAAC 3764
Qy 1097 aGlyThrGlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGly 1117
Db 3765 G-----CTAGAAGTCCCGCCAGTACCAGGAATC----- 3792
Qy 1117 yArgLysLysProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSe 1137
Db 3793 -----CTCCGCTCTCCACCTGTGGA----- 3813
Qy 1137 rAspPheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGly 1157
Db 3814 -----TATCCCCACCC 3824
Qy 1157 uProLeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAl 1177
Db 3825 GCAGTGTCTCTCCGCTGTGCTTCCACACTCTCAGTGGACCGAGTTTCGGAGC 3884
Qy 1177 aAspTrpLeuGlyArgGluLeuGlyGlyCysGln---ProThrThrSerGlyProGluAr 1196
Db 3885 A-----GGAAGAAGTCTAGTCAGTGTGAGGAGGACCACTACCCCAACCACTAT 3935
Qy 1196 gLeu---ThrCysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnPr 1215
Db 3936 GCTGCCCTCATCTCAGCTGTGAGCATTTAGCAGCGAGGAGGACCAACGAGCAACCATCCC 3995
Qy 1215 oSerThrAlaIleLeuGluAla----- 1222
Db 3996 CACACTGTGTGTGACACCACTCACCCACTCGCGAGCTTTGTCTAATCTCTTGTCTACCTCC 4055
Qy 1223 -ProGlyLysSerCysGlnProLysAlaLeuCysProLeuThr-----Va 1237
Db 4056 ACCAATGAGTGAATAGAACCGAAGTC-----CCTTACACACCACTTTTGTCTCAGCC 4109
```

```
Qy 1237 lSerProSerLeuProArgAlaProValSerSerAlaGlnVal 1251
Db 4110 AGGGCCCACTCTTCTTAAGACCACTGTGAAAACACAGCCTCCCTT 4152

RESULT 14
US-09-877-730-23
; Sequence 23, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-23

Alignment Scores:
Pred. No.: 4.35e-51 Length: 1875
Score: 909.00 Matches: 204
Percent Similarity: 52.83% Conservative: 95
Best Local Similarity: 36.04% Mismatches: 205
Query Match: 13.76% Indels: 62
DB: Gaps: 9

US-09-754-997A-2 (1-1252) x US-09-877-730-23 (1-1875)
Qy 470 GlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThrGluGlnVal 489
Db 46 GGTATAAATAATGAAGAGTATCAAGTAGTATCGGAATATGACACAACTCATATTATTATT 105
Qy 490 ArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyrSerGlnLeuGly 509
Db 106 GATCACTTAGAGCTGCGGCAATTTACTTTCATATTGATGATATATCCCAATGGGA 165
Qy 510 AlaSerArgThrSerSerProAlaLeuValHisThrLeuAspValProSerAlaAla 529
Db 166 GCCAGCCAGATGCTGACCATGTGCACACAGAAATCTCTAGAGGATGTTCCCTGAGACCT 225
Qy 530 ProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAlaTrpLeuProLeu 549
Db 226 CTTGAATATTAGTTTGACAAGTGAAGTCCCACTCATATTCTCATCTCTGCTGCAATC 285
Qy 550 ProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGluTyrGlyLeuGlyLys 569
Db 286 CCAGCCAAATATCGCGGGCCAAAGTGTGTGTATCGCTTGTCTTCCGCTTAAGTACT 345
Qy 570 GluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeuThrLeuAsnSer 589
Db 346 GAGAAATCAATCCAAAGTTCGAGCTCCCGGGGACCAAGCATAGTACTTTTGGAGGC 405
Qy 590 LeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGlyValaGlyTyrGly 609
Db 406 CTGAACTGACAGTGTCTACTGTTGCGATTACTGTGCGCCACCAAGAGTGGGCTGGGA 465
Qy 610 ValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGlnSerHisValPro 629
Db 466 GAGTCATCATGTGAGTCTTACATAGGACGCCCAAGCT---ACAAGCGTGAAGGCCCT 522
Qy 630 PheAlaProAlaGluLeuLysValArg---AlaLysMetGluSerLeuValValSerTrp 648
```


QY 505 TyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAsp 524
 Db 415 TATATGCCAATGGGAGCCAGCATGCTGACCATGTGCACACAGATCTAGAGGAT 474
 QY 525 ValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgVal 544
 Db 475 GACCCACAGA----- 483
 QY 545 AlaTyrLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGlu 564
 Db 483 ----- 483
 QY 565 TyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGln 584
 Db 483 ----- 483
 QY 585 LeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThr 604
 Db 483 ----- 483
 QY 605 GlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsn 624
 Db 483 ----- 483
 QY 625 GlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeu 644
 Db 483 ----- 483
 QY 645 ValValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrp 664
 Db 483 ----- 483
 QY 665 GlyGluValGlyThrGluGluGluAlaAspGlyAspArgProProGlyGlyArgGlyAsp 684
 Db 483 ----- 483
 QY 685 GlnAlaTrpAspValGlyProValArgLeuLysLysLysValLysGlnTyrGluLeuThr 704
 Db 483 ----- 483
 QY 705 GlnLeuValProGlyArgProTyrGluValLysLysLeuValAlaPheAsnLysHisGluAsp 724
 Db 484 -----AGAAAATATCATGTGAGACTCTGGCTTACAAACATAGACGAT 528
 QY 725 GlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuProIle 744
 Db 529 GGCATATCAGGCA-----GATCAGACTGTGCAGCACTCCAGGATGCGTGCT 573
 QY 745 GlnArgGly-----ProProLeuProProAlaHisValHisAlaGluSerAsn 760
 Db 574 GTTCGTGATCGCATGGTCCCTCTCCACCACCACCCCATCTCTATGCGAAGGCTAAC 633
 QY 761 SerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleVal 780
 Db 634 ACCTCATCTTCCATCTCTGCATCGGAGAGGCTGCATTCACCGCTGCACAAATCAT 693
 QY 781 AsnTyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThrTyr 800
 Db 694 AACTACACCATCCGCTGTAATCTCTGTGGCTGCAGAAATGCTTCTTGTGTTCTGTACCTT 753
 QY 801 ThrSerSerGlyGluAspIleLeuIleGlyGlyLeuLysProPheThrLysTyrGluPhe 820
 Db 754 CAAACATCAGAACTCACATGTTGTTCAAGTCTAGAACCCAAACACCAATACGAATTT 813
 QY 821 AlaValGlnSerHisGlyValAspMetAspGlyProPhePheGlySerValValGluArgSer 840
 Db 814 GCGCTTCGATTACATGTGATCAGCTTTCACGTCTTGGAGCCCTGTAGTCTTACCATTCT 873
 QY 841 ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
 Db 874 ACTCTTCCAGAGACCCAGCAGGCCCCACCATGTGGAGTAAAGTAGTGCATTAATAGAGGAT 933

QY 861 SerThrValArgLeuHisTrpCysProThrGluProAsnGlyGluIleValGluTyr 880
 Db 934 GACACTGCCCTGGTTCTTGGAAACCCCTGATGGCCAGAAACAGTTGTGACCCGCTAT 993
 QY 881 LeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThr 900
 Db 994 ACTATCTTATATGATCTAGGAGGCTGATTCAGGAGAGTGGCAGGTCTTACACCGT 1053
 QY 901 GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
 Db 1054 GAAGGGGCAATAACCATGGCTTGTCTAGAAAACCTGGTAGCAGGAATGTGTACATTGTC 1113
 QY 921 LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle 940
 Db 1114 AAGATATCTGCATCCAATGAGGTGGAGAGGACCCCTTTTCAAATCTGTGGAGCTGGCA 1173
 QY 941 ThrLeuGlnGluThrPheSerAspSer----- 949
 Db 1174 GTACTTCCAAAGGAAACCTCTGAATCAAATCAGAGGCCCAAGCGTTTAGATTCTGCTGAT 1233
 QY 950 -----LeuAspValHisAlaValThrGlyIleIleVal 960
 Db 1234 GCCAAAGTTTATTTCAGGATATTACCATCTGGACCACAAAATCAATGACTGGCATTTGCTGTA 1293
 QY 961 GlyValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
 Db 1294 GGTGTTGGCATAGCCTTTGACCTGCATCTCATCTGTGTCTCTCATCTTTGATATACCGAAT 1353
 QY 981 SerHisArgGluAla-----LeuProGly 988
 Db 1354 AAAGCCAGGAATCATCTGCTTCCAGACGGCAGAAATGGAACTCAACAGTTTACCTCGT 1413
 QY 989 LeuSerSerSerGlyThrProGlyAsn 997
 Db 1414 ACCAGTGCCTCTTAGCTAGTGGAAAT 1440

Search completed: October 14, 2003, 09:29:47
 Job time : 274 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 15:48:26 ; Search time 10461.9 Seconds
(without alignments)
10933.318 Million cell updates/sec

Title: US-09-754-997A-3

Perfect score: 2796
Sequence: 1 ggggagtgccattgcacca.....cagactccttgatgtgcac 2796

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_em.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2796	100.0	2796	6	AX191261	AX191261 Sequence
2	2796	100.0	3756	6	AX191303	AX191303 Sequence
3	2796	100.0	6176	6	AX191259	AX191259 Sequence
4	2796	100.0	6176	10	AF176694	AF176694 Mus muscu
5	2775.2	99.3	6219	10	AB052820	AB052820 Mus muscu
6	2762.2	98.8	6222	10	AB052821	AB052821 Mus muscu
7	2595.2	92.8	6301	10	AX122535	AX122535 Mus muscu
8	2159.2	77.2	6485	9	AB052622	AB052622 Homo sapi
9	2090.8	74.8	3741	6	AX18406	AX18406 Sequence
10	1491.4	53.3	5549	9	AB046848	AB046848 Homo sapi
C 11	415	14.8	116580	10	AC112162	AC112162 Mus Muscu
C 12	415	14.8	158055	2	AC110235	AC110235 Mus muscu
C 13	376.8	13.5	22427	2	AC106509	AC106509 Rattus no
C 14	331.2	11.8	176282	2	AC068507	AC068507 Homo sapi
15	331.2	11.8	198295	9	AC105129	AC105129 Homo sapi
16	316.8	11.3	2967	9	BC042054	BC042054 Homo sapi
17	316.6	11.3	2976	6	AR237558	AR237558 Sequence
18	316.6	11.3	3210	6	AR237553	AR237553 Sequence
19	316.6	11.3	3219	6	AR237561	AR237561 Sequence
20	316.6	11.3	3874	6	AR237568	AR237568 Sequence
21	316.4	11.3	2700	6	AX497857	AX497857 Sequence
22	315.4	11.3	3453	6	AR237556	AR237556 Sequence
23	302.8	10.8	3198	10	BC053057	BC053057 Mus muscu
24	301.2	10.8	3146	10	AF026465	AF026465 Mus muscu
25	288	10.3	288	6	AX191277	AX191277 Sequence
26	287.4	10.3	2729	9	AK095529	AK095529 Homo sapi
C 27	252.6	9.0	115135	10	AC112161	AC112161 Mus Muscu
28	252	9.0	252	6	AX191281	AX191281 Sequence
29	249	8.9	249	6	AX191273	AX191273 Sequence
30	249	8.9	249	6	AX191275	AX191275 Sequence
31	246	8.8	246	6	AX191279	AX191279 Sequence
32	243	8.7	243	6	AX191265	AX191265 Sequence
33	217	7.8	2481	6	AR237560	AR237560 Sequence
34	217	7.8	2715	6	AR237555	AR237555 Sequence
35	217	7.8	2724	6	AR237562	AR237562 Sequence
36	217	7.8	2958	9	AR237557	AR237557 Sequence
C 37	215.8	7.7	147760	9	AC011846	AC011846 Homo sapi
38	215.8	7.7	176282	2	AC068507	AC068507 Homo sapi
39	213.8	7.6	2139	6	AR237563	AR237563 Sequence
40	213.8	7.6	2382	6	AR237566	AR237566 Sequence
41	195	7.0	195	6	AX191271	AX191271 Sequence
42	192	6.9	192	6	AX191267	AX191267 Sequence
43	189	6.8	189	6	AX191269	AX191269 Sequence
C 44	158.8	5.7	256608	2	AC021040	AC021040 Homo sapi
45	158	5.7	158055	2	AC110235	AC110235 Mus muscu

ALIGNMENTS

RESULT 1	AX191261	Sequence 3 from Patent WO0149714.	2796 bp	DNA	linear	PAT 15-AUG-2001
LOCUS	AX191261	Sequence 3 from Patent WO0149714.				
DEFINITION	AX191261	Sequence 3 from Patent WO0149714.				
ACCESSION	AX191261	Sequence 3 from Patent WO0149714.				
VERSION	AX191261.1	GI:15209512				
KEYWORDS						
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
REFERENCE	Salbaum, J.M.					
AUTHORS	Nope polypeptides, encoding nucleic acids and methods of use					
TITLE	Patent: WO 0149714-A 3 12-JUL-2001;					
JOURNAL						

Neurosciences Research Foundation Inc. (US)

FEATURES

source

Location/Qualifiers

1..2796

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

<1..>2796

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC51121.1"

/db_xref="GI:15209513"

/translation="GELPQRTTTLKSLCDGRLQVILGPQAVVLDCTLGATAGPP

TRVTVKQSDVLEHNLHLLPNSLPLEQSDDEBLRLTWKTEGSLCLAH

SLPVMSQVAVKLTLEDPSLHPESQIVENGTARFECHTKGLPAPILTWKDOVT

VPEERLITLPAWLQILDQSDAGSTRCATNSARQRFQERASLTVALRGSLEATR

GQDVVIAAPENTVSGNVNMBVCASADPTFFVSWVRQDKPISDVIILGRNLL

TASQPRHSVVYCRANPLTRDPATAAEPLRVLAAPISQAPELSTRASFAVC

RASGPRPALHWDGIPLRPNRVRKVGQSGSLVITQIGLDAGYOCVAENSGAT

CAAPLAVVVRGLPSAPTRVATPLSSSSVLVAMERPELHSEQIIGSLHYQKARGV

DNVEYOFANNDITELQVRDLEPNDYEFYVAYSQIGASRTSSPALVHTLDDVPSAA

POLTSSNPSPDIRVAMLPSSLSNGQVLYKIEYGLKEDQVPSFVFNQETLTL

NSLQNKVRISAGTGGAGYGVPSQMOHRTPGVHNSHVFPAPBLKVRKAMESLV

VSWQPPPTQISGLKLYMGVEGTEDRPPGGRDQAWDVGPRVLRKXKQYQEL

TOLVPRPVYKLVAFNKHEDGYAAVWKTEKAPTDLPIQRGPPLPAPVHAESNS

STSILWRKKDPFTTKIVNTVTRFGPMLRNASLVTVYTSGBDILIGLKPPTKYE

FAVQSHGVMDGPFQSVVERSTLPRPSTPDSDLRLSLPSTVPLHWCPTENGEI

VEYLILNNHTQPESHQWTLTTGNIIFSAEVHGLSESDTRIFFMNGATEVGPFPFSR

LQDVITLQETFTSDSLDVH"

BASE COUNT 591 a 849 c 821 g 535 t

ORIGIN

Query Match 100.0%; Score 2796; DB 6; Length 2796;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2796; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGAGCTGCCATTGGCCCGAGAGCAACTGTCAAGCTGAGCTGTGATGAGGACCCCTG 60

Db 1 GGGAGCTGCCATTGGCCCGAGAGCAACTGTCAAGCTGAGCTGTGATGAGGACCCCTG 60

Qy 61 CAAGTGATCTGGGCGCTCAGCAGGCTGTGGTGTGGAGCTGCACTTTGGGGGCTACAGCT 120

Db 61 CAAGTGATCTGGGCGCTCAGCAGGCTGTGGTGTGGAGCTGCACTTTGGGGGCTACAGCT 120

Qy 121 GCTGGGCTCCAGCAGGCTGATGAGCAGAGATGGAGACACTGTACTAGAGCATGAG 180

Db 121 GCTGGGCTCCAGCAGGCTGATGAGCAGAGATGGAGACACTGTACTAGAGCATGAG 180

Qy 181 AACCTGCACCTGCTACCCCAATGGCTCCCTGTGGTGTGGAGCTGCACTTTGGGGGCTACAGCT 240

Db 181 AACCTGCACCTGCTACCCCAATGGCTCCCTGTGGTGTGGAGCTGCACTTTGGGGGCTACAGCT 240

Qy 241 AGCGATGATGAGGAGCTCTTAGGATCTGGAAGTCACTGAGGCGAGCTATTCTGTCTG 300

Db 241 AGCGATGATGAGGAGCTCTTAGGATCTGGAAGTCACTGAGGCGAGCTATTCTGTCTG 300

Qy 301 GCCCAGCCGCTAGGAGTGTGGCGAGCCAGCTGTGGTCAAGCTTGGCAGCTC 360

Db 301 GCCCAGCCGCTAGGAGTGTGGCGAGCCAGCTGTGGTCAAGCTTGGCAGCTC 360

Qy 361 GAAGACTTCTCTGACCCCGAGTCCAGATTTGTGGAGAGAACCGGACAGCAGCTTT 420

Db 361 GAAGACTTCTCTGACCCCGAGTCCAGATTTGTGGAGAGAACCGGACAGCAGCTTT 420

Qy 421 GAATGCCACACCAAGGGCTTCCAGCCCTCATCTTACTTTGGGAAAGGACAGGCTGACC 480

Db 421 GAATGCCACACCAAGGGCTTCCAGCCCTCATCTTACTTTGGGAAAGGACAGGCTGACC 480

Qy 481 GTGCTGAGAGAGCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCTCTAGATGTC 540

Db 481 GTGCTGAGAGAGCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCTCTAGATGTC 540

Qy 541 CAGGACAGTGTGAGGCTCTACCGCTGGTGGCCACCAATTCAGCCCGCCACGATTC 600

Db 541 CAGGACAGTGTGAGGCTCTACCGCTGGTGGCCACCAATTCAGCCCGCCACGATTC 600

541 CAGGACAGTGTGAGGCTCTACCGCTGGTGGCCACCAATTCAGCCCGCCACGATTC 600

601 AGCCAGAGGCGCTCGCTCACTGTGGCGCTCAGAGGCTCTTTGGAGGCTTACCAGGGGGCAG 660

601 AGCCAGAGGCGCTCGCTCACTGTGGCGCTCAGAGGCTCTTTGGAGGCTTACCAGGGGGCAG 660

661 GATGTGTCTATTGTGGCAGCCCGCCAGAGAAACCAACGCGTAGTGTCTGGACAGATGTAGTG 720

661 GATGTGTCTATTGTGGCAGCCCGCCAGAGAAACCAACGCGTAGTGTCTGGACAGATGTAGTG 720

721 ATGAGTGTGGTGGCTCTGTGACCCACCCCTTTTGTGTCTCTGGTCTCGACAGATGGA 780

721 ATGAGTGTGGTGGCTCTGTGACCCACCCCTTTTGTGTCTCTGGTCTCGACAGATGGA 780

781 AAGCTTATCTCCAGGATGTCTGTGCGCGGCGAGCAATCTACTCATCGCCAGCGCG 840

781 AAGCTTATCTCCAGGATGTCTGTGCGCGGCGAGCAATCTACTCATCGCCAGCGCG 840

841 CAGCTCTGGCACTCTGGAGTCTATGTCTGCGGAGCCCAACAGCCCTCTCACGCGTACTTC 900

841 CAGCTCTGGCACTCTGGAGTCTATGTCTGCGGAGCCCAACAGCCCTCTCACGCGTACTTC 900

901 GCCACTGCGCTGTGAGCTCTCGAGTGTCTGCTGCGCCAGGCATCTCGCAGGACCCGAG 960

901 GCCACTGCGCTGTGAGCTCTCGAGTGTCTGCTGCGCCAGGCATCTCGCAGGACCCGAG 960

961 GCGCTCTGCGGAGCGCGGCGCAGCAGCGCGCTTGTGTGCGCGGCGCTCGCGGGAGCCA 1020

961 GCGCTCTGCGGAGCGCGGCGCAGCAGCGCGCTTGTGTGCGCGGCGCTCGCGGGAGCCA 1020

1021 CGGCGCGCTGCACTGTGCTGCAAGCGGATCCGCTTGGCAGCCCAATGGGCGCTCAAG 1080

1021 CGGCGCGCTGCACTGTGCTGCAAGCGGATCCGCTTGGCAGCCCAATGGGCGCTCAAG 1080

1081 GTGACGCGCTGTGCGGCGCTGTGCTCATCTCAGATGCGCTCTGACGAGCGCTGGCTAC 1140

1081 GTGACGCGCTGTGCGGCGCTGTGCTCATCTCAGATGCGCTCTGACGAGCGCTGGCTAC 1140

1141 TACCATGCTGTAGCAGAAACAGCGCGGAGACTGCTGTGCGCTGCGCCCTCGCGGTA 1200

1141 TACCATGCTGTAGCAGAAACAGCGCGGAGACTGCTGTGCGCTGCGCCCTCGCGGTA 1200

1201 GTGTGCGCGAGGCGCTGCGCCAGCGCCCGCTCGGGTCAAGCCAGCCCGCTGAGCAGC 1260

1201 GTGTGCGCGAGGCGCTGCGCCAGCGCCCGCTCGGGTCAAGCCAGCCCGCTGAGCAGC 1260

1261 TCCTCTGTGCTGTGGCGCTGTGGAGCGGCTTGAGTTTGCAAGCGCAAAATCATTTGGCTTC 1320

1261 TCCTCTGTGCTGTGGCGCTGTGGAGCGGCTTGAGTTTGCAAGCGCAAAATCATTTGGCTTC 1320

1321 TCTCTTCACTACCAAGGAGGAGGAGTGGACATGTGGAGTACAGTTTTCAGTAAAC 1380

1321 TCTCTTCACTACCAAGGAGGAGGAGTGGACATGTGGAGTACAGTTTTCAGTAAAC 1380

1381 AATGACACACAGAGCTGCGAGTTCGGACCTTGGAAACCAACACGGATTTATGAGTTCTAC 1440

1381 AATGACACACAGAGCTGCGAGTTCGGACCTTGGAAACCAACACGGATTTATGAGTTCTAC 1440

1441 GTGTGGCTTACTCTCAGCTGTGGGCGCAGCCGAACTTCAGCCAGCCCTGGTGCATACA 1500

1441 GTGTGGCTTACTCTCAGCTGTGGGCGCAGCCGAACTTCAGCCAGCCCTGGTGCATACA 1500

1501 CTGGAAGATGTCTCCAGCGAGCAGCCAGCTTACTTGTCTGAGCCCGCCAGCCCTCGGAC 1560

1501 CTGGAAGATGTCTCCAGCGAGCAGCCAGCTTACTTGTCTGAGCCCGCCAGCCCTCGGAC 1560

1561 ATCAGGCTGGCATGTGCTGCTCCCTGCTCCAGCTTGGAGCAATGGACAGTGTGAGTAC 1620

1561 ATCAGGCTGGCATGTGCTGCTCCCTGCTCCAGCTTGGAGCAATGGACAGTGTGAGTAC 1620

1621 AAGATAGAGTACCGTTTGGGGAGGAAGATCAGGTTTTCTCCACCGAGGCTGCTGGAAT 1680

1621 AAGATAGAGTACCGTTTGGGGAGGAAGATCAGGTTTTCTCCACCGAGGCTGCTGGAAT 1680

[illegible]

1321 TCTCTTCACTACCAAAAGGCAAGGGAGTGGACAATGTGGAGTACCAAGTTTGCAGTAAC 1380
1384 TCTCTTCACTACCAAAAGGCAAGGGAGTGGACAATGTGGAGTACCAAGTTTGCAGTAAC 1443
1381 AATGACACACAGAGCTGAGGTTGGGACCTGGAAACCCAAACACAGGATTAATGAGTTCTAC 1440
1444 AATGACACACAGAGCTGAGGTTGGGACCTGGAAACCCAAACACAGGATTAATGAGTTCTAC 1503
1441 GTGGTGGCCCTACTCCAGCTGGGGGCGAGCCGAACTCCAGAGCCAGCCCTGGTGGATACA 1500
1504 GTGGTGGCCCTACTCCAGCTGGGGGCGAGCCGAACTCCAGAGCCAGCCCTGGTGGATACA 1563
1501 CTGGAGCATGTCTCCAGCCGAGCACCCAGCTTACCTTGTCCAGCCCAACCCCTCGGAC 1560
1564 CTGGAGCATGTCTCCAGCCGAGCACCCAGCTTACCTTGTCCAGCCCAACCCCTCGGAC 1623
1561 ATCAGGCTGGATGGTCCCTGCTCCAGCTGAGCAATGAGCAAGTGTGAGTAC 1620
1624 ATCAGGCTGGATGGTCCCTGCTCCAGCTGAGCAATGAGCAAGTGTGAGTAC 1683
1621 AAGATAGACTAGGTTTGGGAGGAAGATCAGGTTTCTCCAGAGGTGCTGGAAT 1680
1684 AAGATAGACTAGGTTTGGGAGGAAGATCAGGTTTCTCCAGAGGTGCTGGAAT 1743
1681 GAGACAACTTACGTTAACTCCTTACGACCAAAACAAAGTGTACCGAGTTCGGATTCA 1740
1744 GAGACAACTTACGTTAACTCCTTACGACCAAAACAAAGTGTACCGAGTTCGGATTCA 1803
1741 GCTGGCACTGGGCTGGCTATGAGTCCCTTCTCAGTGATGACAGACACCTGGT 1800
1804 GCTGGCACTGGGCTGGCTATGAGTCCCTTCTCAGTGATGACAGACACCTGGT 1863
1801 GTGCACAAACAGAGCATGTTCCCTTGGCCCTGACAGATTCAGGTGAGGCGCAAGATG 1860
1864 GTGCACAAACAGAGCATGTTCCCTTGGCCCTGACAGATTCAGGTGAGGCGCAAGATG 1923
1861 GAGTCCCTGGTGTGTATGAGGAGCGCCCTTCCACCCACAGATCTCTGGATACAAA 1920
1924 GAGTCCCTGGTGTGTATGAGGAGCGCCCTTCCACCCACAGATCTCTGGATACAAA 1983
1921 CTCTACTGGGAGAGTGGGAAACAGAGGAGGAGGAGTGTGACCGCCCCCGAGGGGT 1980
1984 CTCTACTGGGAGAGTGGGAAACAGAGGAGGAGGAGTGTGACCGCCCCCGAGGGGT 2043
1981 CGTGGAGATCAAGCTTGGGACGTCGGGCGCGTGGCTGAGAGAGAGTGAAGCAGTAT 2040
2044 CGTGGAGATCAAGCTTGGGACGTCGGGCGCGTGGCTGAGAGAGAGTGAAGCAGTAT 2103
2041 GAACTGACCCAGTTAGTCTCTGGCAGGCGCTACGAGTGAAGTCTGTAGCTTTCAACAAA 2100
2104 GAACTGACCCAGTTAGTCTCTGGCAGGCGCTACGAGTGAAGTCTGTAGCTTTCAACAAA 2163
2101 CACGAGGACGGCTACGCTGTGTGGAAAGGCAAGACGAGAGGCGCCCAAGCCAGAC 2160
2164 CACGAGGACGGCTACGCTGTGTGGAAAGGCAAGACGAGAGGCGCCCAAGCCAGAC 2223
2161 CTGCTCTACGAGGGGGGACCGCTGCTCTGCTCCATGTCCAGGAGTCAAGGATCAACAGC 2220
2224 CTGCTCTACGAGGGGGGACCGCTGCTCTGCTCCATGTCCAGGAGTCAACAGC 2283
2221 TCCACTTCCATTTGGCTTGGTGGAAAGAGCCAGACTTTACCACTGTCAAGATTGTCAAC 2280
2284 TCCACTTCCATTTGGCTTGGTGGAAAGAGCCAGACTTTACCACTGTCAAGATTGTCAAC 2343
2281 TACACTGTACGCTTCGGCCCTGGGGCTCAGGAATGCTTCCCTGGTCACTACTATACC 2340
2344 TACACTGTACGCTTCGGCCCTGGGGCTCAGGAATGCTTCCCTGGTCACTACTATACC 2403
2341 AGCTCTGGAGAGACATTTCTCATTTGGGGCTGAAACCATTTACCAAGTACGATTGGG 2400
2404 AGCTCTGGAGAGACATTTCTCATTTGGGGCTGAAACCATTTACCAAGTACGATTGGG 2463
2401 GTACAGTCCCAACGAGTGGATATGGATGGGCGCCCTTTGGCTCGCTGAGAACGCTCCACC 2460

2464 GTACAGTCCCAACGAGTGGATATGGATGGGCGCTTTGGCTCCGTAGAACGCTCCACC 2523
2461 CTGCTGACCGGCTTCAACACCTCTCTGACCTGGCGCTGAGCCCTGACACCATCC 2520
2524 CTGCTGACCGGCTTCAACACCTCTCTGACCTGGCGCTGAGCCCTGACACCATCC 2583
2521 ACCGTTCCGTTTACACTGGTGTCCCGCCACGAGGCGCCAAATGGTGGAGTGTGATCTA 2580
2584 ACCGTTCCGTTTACACTGGTGTCCCGCCACGAGGCGCCAAATGGTGGAGTGTGATCTA 2643
2581 ATTCTCTACAGAACCAACCAACCCAGGCGCCAAACACAGTGGACACTCTCAACAGAG 2640
2644 ATTCTCTACAGAACCAACCAACCCAGGCGCCAAACACAGTGGACACTCTCAACAGAG 2703
2641 GGAAACATCTTTCAGTGCAGAGTCCATGGCTTAGAGTGCACCTCGGTATTTCTTCAAG 2700
2704 GGAAACATCTTTCAGTGCAGAGTCCATGGCTTAGAGTGCACCTCGGTATTTCTTCAAG 2763
2701 ATGGAGGCGCCGACAGAGTGGGCGCTTGGGCGCTTTTCCCGCTTGCAGGATGTGATTACT 2760
2764 ATGGAGGCGCCGACAGAGTGGGCGCTTGGGCGCTTTTCCCGCTTGCAGGATGTGATTACT 2823
2761 CTGCAAGAGACATTTCTCAGACTCTCTTGGATGTGCAC 2796
2824 CTGCAAGAGACATTTCTCAGACTCTCTTGGATGTGCAC 2859

RESULT 4
AF176694
LOCUS
DEFINITION Mus musculus neighbor of Punc e11 protein (Nope) mRNA, complete cds.
ACCESSION AF176694
VERSION AF176694.1 GI:7650185
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Salbaum, J.M. and Kappen, C.
TITLE Cloning and expression of nope, a new mouse gene of the immunoglobulin superfamily related to guidance receptors
JOURNAL Genomics 64 (1), 15-23 (2000)
MEDLINE 20175427
PUBMED 10708514

REFERENCE
AUTHORS Salbaum, J.M.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1999) The Neurosciences Institute, 10640 John Jay
Hopkins Drive, San Diego, CA 92121, USA

FEATURES
source
1..6176
Location/Qualifiers

1..6176
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB"
/db_xref="taxon:10090"
/chromosome="9"
/map="band D-E1; 24.59 cR from D9Mit143"
1..6176
/gene="Nope"
1..3759
/gene="Nope"
/note="immunoglobulin superfamily; related to guidance receptors"

gene
CDS
1..3759
/product="neighbor of Punc e11 protein"
/protein_id="AAF65930.1"
/db_xref="GI:7650186"

1..3759
/translation="MARADPTGRVTVMSKDGDTVLEHNLHLLPNSLWLSPLQEDS
LCPEQAVVLDCTLGATAAGAPTRVTVMSKDGDTVLEHNLHLLPNSLWLSPLQEDS
DDEALRIWKVTEGSGYSLAHSPLGWASQAVAVKLTLEDPLHPSQIVBENGTA

PECHTKGLPAPIITWEXDQVTVPEEPRILITLPRKWLQILQVQSDAGSYRCVATNSAR
ORFSQELSTVALRGLSEATRGQDVIIVAAPENTTVVSGQNVNVECVASADPTFPVSW
VROQKISTDVI VLGRITNLLIASAQPHRSQVYCRANKPLTRDFAFAAABLRVLAAP
AISQAPRALSRTRASTARFVCRASGEPRPALHMDGI PLRPNRGRVQVQGGSLVIT
OIGLQDAGYQCVALENSAGTACAAPLANVVRREGLPSPATRVATPLSSSLVAMER
PLHSEBQIGSLHYQKARGVDNVEYQFVANDTTELQVRLDLEPNTDYEFPVAYSOL
GASTSPALVHTLDDVPSAPOOLTSSPNPSDIRVAMLPLPSSLSNGVLKRYEYGL
LGKEDQVFEVPEQNETQLTNSIQPNKXVVRVIRISAGTGAGYGVPSQMOHRTPGVHN
QSHVPPAPBELKVRKAKMESLIVSWQPPPHPTQISGYKLWYGEVGTBEADGDRPQGR
GDOAWDVPRLKXVQYELTQVPCRPVEVKLVA FNKHEDEGVAA WNKGTREKAPTV
DLPIORGPPPLPPHVAHSNSTSISIMLRWKKPDPFTTKLYNYTVRFPWGLRPNLSLTV
YYSQSGDILLIGLHAFKYEFAVQSHGVDMDGFPFVSVERSTLPLRPSDLRLUS
PLTSTVRLHWCPTPEPGEIYELIILYSNNHTOPEHWTLLTTEGNI FSAEHLGLS
DTRYFFKMGARTYEGPGRFRLQDVIITLQETFSDSLVDHVAVTGII VGVCLGLCLLAC
MCDVRSQSHREALPLGSSQTPGNPALYTRRAQPPSPVAAHELSLVHVRPQDWPSP
PPSDVEKAEVHSLWAGSVDRCRHSKRKISWAQAGPNWAGSWAGCELPOGSGPRPA
LTRALLPAGTGQTLILQALVYDGI KSNGRKKPSPACRNOVEARVIVHSDFQASKGCP
DLHQDLPEBPRLTAETLPSSTGAVDIUSQGMADWIGRLBGLGCOPTTSPERNTCLPEAA
SASCSDLPQLPSTAIIBRAPGKSCQPKALCPLTVSPSLPRAPVSSAQPV"

misc_feature

1..23

/gene="Nope"

/note="unable to sequence region in mRNA"

BASE COUNT 1316 a 1756 c 1634 g 1470 t

ORIGIN

Query Match 100.0%; Score 2796; DB 10; Length 6176;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2796; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGGAGCTGC	CAATTTGCC	CCAGGAGCA	CACTCTCAAGCTG	AGCTGTGATG	AGGGACCCCTG	60
Db	64	GGGAGCTGC	CAATTTGCC	CCAGGAGCA	CACTCTCAAGCTG	AGCTGTGATG	AGGGACCCCTG	123
Qy	61	CAAGTGATC	TCTGGCC	CTGAGCAG	CGCTGTGTCTG	CTGACTG	CTTTGGGGCTAC	120
Db	124	CAAGTGATC	TCTGGCC	CTGAGCAG	CGCTGTGTCTG	CTGACTG	CTTTGGGGCTAC	183
Qy	121	GCTGGGCT	CCGAC	CAGGGTGA	CTGAGCAAG	AGAGTGA	CTGTACTAGACATG	180
Db	184	GCTGGGCT	CCGAC	CAGGGTGA	CTGAGCAAG	AGAGTGA	CTGTACTAGACATG	243
Qy	181	AACCTGCA	CTCTG	TACCAAT	TGGCTTCC	CTGTGCTC	TCCACCCCTAGAGCAAGAGAC	240
Db	244	AACCTGCA	CTCTG	TACCAAT	TGGCTTCC	CTGTGCTC	TCCACCCCTAGAGCAAGAGAC	303
Qy	241	AGCGATGAT	GAGGAAG	CTCTTAG	ATCTGGAAG	GTCTGAG	AGGCGAGCTATTCCTGTCTG	300
Db	304	AGCGATGAT	GAGGAAG	CTCTTAG	ATCTGGAAG	GTCTGAG	AGGCGAGCTATTCCTGTCTG	363
Qy	301	GCCACAC	CGCCCT	AGAGTGT	GTGCC	CAGCAG	GTGTGTGTCTGCTCAAGCTTTGCCACATC	360
Db	364	GCCACAC	CGCCCT	AGAGTGT	GTGCC	CAGCAG	GTGTGTGTCTGCTCAAGCTTTGCCACATC	423
Qy	361	GAGACTTCT	CTCTG	CAACCC	CGAGTCC	CAGATTGT	GAGGAGAAACGGGACAGCAGCTTT	420
Db	424	GAGACTTCT	CTCTG	CAACCC	CGAGTCC	CAGATTGT	GAGGAGAAACGGGACAGCAGCTTT	483
Qy	421	GAATGCC	CAACCA	AGGGCTT	CCAGCC	CCCATCA	TACTTGGGAAAAGGACCAAGTGACC	480
Db	484	GAATGCC	CAACCA	AGGGCTT	CCAGCC	CCCATCA	TACTTGGGAAAAGGACCAAGTGACC	543
Qy	481	GTGCTG	GAGGAG	CCCGGCT	CATCACT	CTTCC	CAAGTGGCTCTCCAGATCCTAGATGTC	540
Db	544	GTGCTG	GAGGAG	CCCGGCT	CATCACT	CTTCC	CAAGTGGCTCTCCAGATCCTAGATGTC	603
Qy	541	CAGGAC	AGTGAT	GATGAG	GTCTT	ACCGT	GTGGCCACCAATTCAGCCCGCCCAACGATTC	600
Db	604	CAGGAC	AGTGAT	GATGAG	GTCTT	ACCGT	GTGGCCACCAATTCAGCCCGCCCAACGATTC	663
Qy	601	AGCCAG	GAGGCT	CGCTCA	CTGTG	GGCCCT	CAGAGGGCTTTTGGAGGCTTACAGGGGGCAG	660
Db	664	AGCCAG	GAGGCT	CGCTCA	CTGTG	GGCCCT	CAGAGGGCTTTTGGAGGCTTACAGGGGGCAG	723

```
Db 1804 GCTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGGATGACGACGACGACCTGGT 1863
Qy 1801 GTGGCAACACGAGCATGTTCCCTTGGCCCTTCGAGATTTGAAGGTGAGGGCAAGATG 1860
Db 1864 GTGGCAACACGAGCATGTTCCCTTGGCCCTTCGAGATTTGAAGGTGAGGGCAAGATG 1923
Qy 1861 GAGTCCCTGGTGGTGTCTAGTGGCAGCGCCCTTCACCCACCCAGATCTCTGGATACAAA 1920
Db 1924 GAGTCCCTGGTGGTGTCTAGTGGCAGCGCCCTTCACCCACCCAGATCTCTGGATACAAA 1983
Qy 1921 CTCTACTGGGGAGAGGTGGAAACAGAGGAGGAGGAGATGTGTGACCGCCCCCAGGGGT 1980
Db 1984 CTCTACTGGGGAGAGGTGGAAACAGAGGAGGAGGAGATGTGTGACCGCCCCCAGGGGT 2043
Qy 1981 CGTGGAGATCAAGCTTGGGACCTCGGGCCCGTGGCTGGAAGAAAGTGAAGCAGTAT 2040
Db 2044 CGTGGAGATCAAGCTTGGGACCTCGGGCCCGTGGCTGGAAGAAAGTGAAGCAGTAT 2103
Qy 2041 GAACTGACCCAGTTAGTCCCTGGCAGGCCGTACGAGGTGAAGCTGTTAGCTTTCAACAAA 2100
Db 2104 GAACTGACCCAGTTAGTCCCTGGCAGGCCGTACGAGGTGAAGCTGTTAGCTTTCAACAAA 2163
Qy 2101 CACGAGGACGGTACGCTCTGTGTGAAAGGCAAGACGAGAGCGCCACGCGAC 2160
Db 2164 CACGAGGACGGTACGCTCTGTGTGAAAGGCAAGACGAGAGCGCCACGCGAC 2223
Qy 2161 CTGCCCTATCAGAGGGGGCCACGCTGCTCTCTGCCATGTGTCAACGAGGTCAACACGC 2220
Db 2224 CTGCCCTATCAGAGGGGGCCACGCTGCTCTCTGCCATGTGTCAACGAGGTCAACACGC 2283
Qy 2221 TCCACTTCCATTTGGCTTGGTGGAGAGCAGCAGCTTTACCACTGTCAAGATTGTCAAC 2280
Db 2284 TCCACTTCCATTTGGCTTGGTGGAGAGCAGCAGCTTTACCACTGTCAAGATTGTCAAC 2343
Qy 2281 TACACTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCCTGTGTCACTACTATACC 2340
Db 2344 TACACTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCCTGTGTCACTACTATACC 2403
Qy 2341 AGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACAGATGACAGTTTGG 2400
Db 2404 AGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACGAGTTTGG 2463
Qy 2401 GTACAGTCCACGAGTGGATATGATGGCGCTTGGCTCGTGTGTAGAACGCTCCACC 2460
Db 2464 GTACAGTCCACGAGTGGATATGATGGCGCTTGGCTCGTGTGTAGAACGCTCCACC 2523
Qy 2461 CTGCCCTGACCGGCTTTCAACACCTCTCTGTGACCTGCGCTGAGCCCTCGACCATCC 2520
Db 2524 CTGCCCTGACCGGCTTTCAACACCTCTCTGTGACCTGCGCTGAGCCCTCGACCATCC 2583
Qy 2521 ACCGTTCCGTTTACATGGTGTCCCTCCACGAGGCCAATGGTGGAGATTGTGGAGTATCTA 2580
Db 2584 ACCGTTCCGTTTACATGGTGTCCCTCCACGAGGCCAATGGTGGAGATTGTGGAGTATCTA 2643
Qy 2581 ATTCTCTACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAG 2640
Db 2644 ATTCTCTACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAG 2703
Qy 2641 GGAACCATCTTACGTGACAGAGTTCATGGCTTAGAGAGTGACACTCGGTATTTCTTCAAG 2700
Db 2704 GGAACCATCTTACGTGACAGAGTTCATGGCTTAGAGAGTGACACTCGGTATTTCTTCAAG 2763
Qy 2701 ATGGGAGCCGCAACAGAGTGGGGCTTGGGCCCTTTTCCGCTTTCGAGATGTGATTCT 2760
Db 2764 ATGGGAGCCGCAACAGAGTGGGGCTTGGGCCCTTTTCCGCTTTCGAGATGTGATTCT 2823
Qy 2761 CTGCAAGACATCTTCAAGTCTCTTGGATGTGCAC 2796
Db 2824 CTGCAAGACATCTTCAAGTCTCTTGGATGTGCAC 2859
```

RESULT 5

AB052620
LOCUS
DEFINITION Mus musculus DDM36 mRNA, complete cds.
ACCESSION AB052620
VERSION AB052620.1 GI:11862938
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source

AB052620 6219 bp mRNA linear ROD 14-APR-2001
Mus musculus DDM36 mRNA, complete cds.
AB052620.1 GI:11862938
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Murakami,H., Nakamata,T., Nakayama,T., Yamamoto,H., Hosaka,T.,
Aoyama,T., Nagayama,S., Oka,M., Kiyono,T., Sasaki,M.S., Nakamura,T.
and Toguchida,J.
Up-regulation of a ras effector and down-regulation of a cell
adhesion molecule are associated with transformation of osteoblasts
Unpublished
2 (bases 1 to 6219)
Toguchida,J., Nakamata,T., Murakami,H., Nakayama,T. and Nakamura,T.
Submitted (11-DEC-2000) Junya Toguchida, Kyoto University,
Institute for Frontier Medical Sciences, 53 Kawahara-cho, Shogoin,
Sakyo-ku, Kyoto, Kyoto 606-8507, Japan
(E-mail:togjun@frontier.kyoto-u.ac.jp, Tel:81-75-751-4134,
Fax:81-75-751-4144)
Location/Qualifiers
1. 6219
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6BA"
/db_xref="taxon:10090"
/clone="DDM36"
/cell_line="MMC2"
/cell_type="osteoblast"
/note="similar to Nope (NM_020043)-vector:pDDM36ORP"

gene

CDS

1. 6219
/gene="DDM36"
50..3808
/gene="DDM36"
/note="similar to Nope (NP_064427)"
/codon_start=1
/evidence=experimental
/product="DDM36"
/protein_id="BAB19278.1"
/db_xref="GI:11862938"
/translation="MARADTGRGLLVLPCLLSARGELPLPQETTVKLSCEGGLPQVI
LGPOAVVLDCTGATAGPRVTVMSKGDVTVLEHNLHLLPNSLSPSQBDS
DEBALRIWKVTGYSCLAHSPLGWSQVAVVVKLATLEDFSLHPSQIYEENGTAR
PECHTKGPAPIITWKKQVTVBESRLITLPNGVLQILDVQSDAGSYRCVNTSAR
QRFSOASLTVALRGSLEATRGQDVIIVAAPEVTVVSGQSVWMEVASADPTFVSW
VRQGPISITDVIVLGRNLLISAQPHSGVYVCRANKPRTPRPATAAELVLAAP
AISQAPALSRSTRASTARPVCRASGEPFALHLDGIPLPNGRVKVGCGGSLVIT
QIGLDAGYQCVASNSAGTACAAAPLAVVVRGLPSPATRVATPFLSSSVLVAMER
PELHSEQIIGFSLHYKARGVDNVEYQFVANNDDTTELVQRLDLEPTVEFTVVAISQL
GASRTSSPALVHTLDDVPAAAPQLTASPNPSDIRVAMPLPSSLNSQVLRKYLEYG
LGKDDVPSTVRPNETQLTNSLPKNKXVVRVRSAGTGVAGVGPVQWQVHRTQVHN
QSHVPAPAEIKYAKMESLVVSHQPPPHPTQISGYKLYWRBVGTEEADGDRPPGR
GDQWVGVFVKLKKVQELTQLVPGRLYEVKLVAFNKHSDGTAAVWNGKRLKAPT
DLPTGPPPLPAPHVAESNSTSIWLRKKPDPFTTVKIVNYTVVTFPGWGLRNASLV
YVTSGEDILIGLKPPTKYRFVAVQSHGDMDPFGSVVVERSTLPDRSTPSPDLRLS
PLTPVRLHWCPTTPNGEIVEYVILYNNHTOPHQWTLTATBGNIFSAEVLHGLS
DTRVFPKMGATEVGRPGSRLODVTIQTQPSDLDVHVTGIIIVGVLGILCLLAC
MCAGLRSSHREALPQLSSGTPGNPALYTRARLPSPVPAHAELSLVHRPQWSP
PPSDVEDKAVHSLMGGSVSDCRHSKRRKISNAQGGPNWAGSWAGSLPFGSGKCP
LTRALLPAGTQTLQLQALVDAIKSNGRKKPSACRQVBEAVIYHSDFGASGCP
DLHLQDLEPBLTAETLPTSGAVDSLQGDWLRGLGCGQPTTSGPERLTCLPEAA
SASCSCLQPSLTAEAPGKSCQKALCPLTVSPSLPRAVPSSAQVP"

BASE COUNT 1319 a 1772 c 1652 g 1476 t
ORIGIN

Query Match 99.3%; Score 2775.2; DB 10; Length 6219;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2783; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGAGCTGCCATTGGCCCGAGGAGCAAACTGTCTAAGCTGTAGCTGTGTGATGAGGGACCCCTG 60
Db 113 GGGAGCTGCCATTGGCCCGAGGAGCAAACTGTCTAAGCTGTAGCTGTGTGATGAGGGACCCCTG 172
Qy 61 CAAGTATCTTGGGCCCTGAGCAGGCTGTGTGCTGTGCACTGTGCTGTTGGGGGCTTACAGCT 120
Db 173 CAAGTATCTTGGGCCCTGAGCAGGCTGTGTGCTGTGCACTGTGCTGTTGGGGGCTTACAGCT 232
Qy 121 GCTGGGCTCCGACAGGGTGTACATGAGCAGGATGAGACACATGTACTAGAGCATGAG 180
Db 233 GCTGGGCTCCGACAGGGTGTACATGAGCAGGATGAGACACATGTACTAGAGCATGAG 292
Qy 181 AACCTGACCTGTCTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAGAC 240
Db 293 AACCTGACCTGTCTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAGAC 352
Qy 241 AGCGATGATGAGGAGCTCTTAGATCTGAGAGGTCTCTAGGGGACGCTATTCCTGTCTG 300
Db 353 AGCGATGATGAGGAGCTCTTAGATCTGAGAGGTCTCTAGGGGACGCTATTCCTGTCTG 412
Qy 301 GCCCAGCCGCTAGGAGTGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGGCCACACTC 360
Db 413 GCCCAGCCGCTAGGAGTGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGGCCACACTC 472
Qy 361 GAAGACTTCTCTCTGCAACCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGCTTT 420
Db 473 GAAGACTTCTCTCTGCAACCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGCTTT 532
Qy 421 GAATGCCACCAAGGGCTTCCAGCCCCCATCATTTCTTGGGAAAGGACCAAGTGAAC 480
Db 533 GAATGCCACCAAGGGCTTCCAGCCCCCATCATTTCTTGGGAAAGGACCAAGTGAAC 592
Qy 481 GTGCCTGAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCCTAGATGTC 540
Db 593 GTGCCTGAGGAGTCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCCTAGATGTC 652
Qy 541 CAGGACAGTATGATGAGGCTCTTACCGCTGTGGCCACCAATTCAAGCCCGCCACAGATTC 600
Db 653 CAGGACAGTATGATGAGGCTCTTACCGCTGTGGCCACCAATTCAAGCCCGCCACAGATTC 712
Qy 601 AGCCAGAGGCTCGCTCATCTGTGGCCCTCAGAGGGTCTTGTGAGGCTTACAGGGGGCAG 660
Db 713 AGCCAGAGGCTCGCTCATCTGTGGCCCTCAGAGGGTCTTGTGAGGCTTACAGGGGGCAG 772
Qy 661 GATGTGTCTATTGGCAGCCCAAGACACACACAGTGTGTCTGACAGATGTAGTG 720
Db 773 GATGTGTCTATTGGCAGCCCAAGACACACACAGTGTGTCTGACAGATGTAGTG 832
Qy 721 ATGAGTGTGCTGGCTCTGTCTGACCCACCCCTTTTGTCTGGTCCGACAGATGGA 780
Db 833 ATGAGTGTGCTGGCTCTGTCTGACCCACCCCTTTTGTCTGGTCCGACAGATGGA 892
Qy 781 AAGCCTATCTCCAGGATGTATGTCTGGCCCGGACCAATCTACTCATCGCCAGCGG 840
Db 893 AAGCCTATCTCCAGGATGTATGTCTGGCCCGGACCAATCTACTCATCGCCAGCGG 952
Qy 841 CAGCCTCGGCACTGTGAGTCTATGTCTGCGAGCCAAAGAGCCCTCAAGCGGTGACTTC 900
Db 953 CAGCCTCGGCACTGTGAGTCTATGTCTGCGAGCCAAAGAGCCCTCAAGCGGTGACTTC 1012
Qy 901 GCCATCGGCTGTGAGTCTCGAGTCTTGTCTGCCCGGACCATCTCGAGGACCCGAG 960
Db 1013 GCCATCGGCTGTGAGTCTCGAGTCTTGTCTGCCCGGACCATCTCGAGGACCCGAG 1072
Qy 961 GCGCTCTCGGCGACGCGGGCCAGCAGCGGCTTCTGTGTGCTGGGGGCTTCGGGGAGCCA 1020
Db 1073 GCGCTCTCGGCGACGCGGGCCAGCAGCGGCTTCTGTGTGCTGGGGGCTTCGGGGAGCCA 1132
Qy 1021 CGGCCCGGCTGTGCTGTGCTGACAGAGGATCCCGTTGTGGACCCCAATGGGCGGCTCAAG 1080
Db 1133 CGGCCCGGCTGTGCTGTGCTGACAGAGGATCCCGTTGTGGACCCCAATGGGCGGCTCAAG 1192

Qy 1081 GTGCAAGGCGGCTGGCGGCAAGCTTGTGTCTCATCTCAAGATCGGCTGTGCAAGGACGCTGTGCTAC 1140
Db 1193 GTGCAAGGCGGCTGGCGGCAAGCTTGTGTCTCATCTCAAGATCGGCTGTGCAAGGACGCTGTGCTAC 1252
Qy 1141 TACCACTGTGTAGCAGAAACACGCGGGGAACTGCTCTGTGTGCTGCGGCTTGGCGGCTA 1200
Db 1253 TACCACTGTGTAGCAGAAACACGCGGGGAACTGCTCTGTGTGCTGCGGCTTGGCGGCTA 1312
Qy 1201 GTGGTGTGCGAGGGGCTGCGCAGCGCCGACCTCGGCTCACAGCACGCGCTGTGAGCAGC 1260
Db 1313 GTGGTGTGCGAGGGGCTGCGCAGCGCCGACCTCGGCTCACAGCACGCGCTGTGAGCAGC 1372
Qy 1261 TCTCTGTGTGTGTGCTGTGGAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1320
Db 1373 TCTCTGTGTGTGTGCTGTGGAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1432
Qy 1321 TCTCTTCACTTACCAAAAGGACGAGGGGAGTGGACAAATGTGTGAGTACCAATTTGTGCAATTAAC 1380
Db 1433 TCTCTTCACTTACCAAAAGGACGAGGGGAGTGGACAAATGTGTGAGTACCAATTTGTGCAATTAAC 1492
Qy 1381 AATGACACACAGAGCTGCAAGGTTCTGGGACCTCGAACCCCAACAGGATTTATGATTTCTAC 1440
Db 1493 AATGACACACAGAGCTGCAAGGTTCTGGGACCTCGAACCCCAACAGGATTTATGATTTCTAC 1552
Qy 1441 GTGGTGTGCTTACTCCAGCTGTGGGGCCAGCCGAACTCTCAGCCAGCCCTGTGTGCTATCA 1500
Db 1553 GTGGTGTGCTTACTCCAGCTGTGGGGCCAGCCGAACTCTCAGCCAGCCCTGTGTGCTATCA 1612
Qy 1501 CTGGAAGATGTCTCCAGCGGAGCAGCCAGCTTACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db 1613 CTGGAAGATGTCTCCAGCGGAGCAGCCAGCTTACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1672
Qy 1561 ATCAGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Db 1673 ATCAGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1732
Qy 1621 AAGATAGATGACGTTTGGGGAAAGGAGATCAGGTTTCTTCTCAACGAGGTCCTGTGAAAT 1680
Db 1733 AAGATAGATGACGTTTGGGGAAAGGAGATCAGGTTTCTTCTCAACGAGGTCCTGTGAAAT 1792
Qy 1681 GAGACACAACTTACGTTTAAACTCATCTTCAAGCCAAAACAAAGTGTACCGAGTCCGAGTTTCA 1740
Db 1793 GAGACACAACTTACGTTTAAACTCATCTTCAAGCCAAAACAAAGTGTACCGAGTCCGAGTTTCA 1852
Qy 1741 GCTGGCACTGGCGCTGCTGTATGAGTCTTCTTCTCAGTGTGATGAGCAGACACGCTGCT 1800
Db 1853 GCTGGCACTGGCGCTGCTGTATGAGTCTTCTTCTCAGTGTGATGAGCAGACACGCTGCT 1912
Qy 1801 GTGCACAAACAGAGCCATGTTTCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1913 GTGCACAAACAGAGCCATGTTTCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1972
Qy 1861 GAGTCCCTGTGTGTCTATGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1973 GAGTCCCTGTGTGTCTATGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2032
Qy 1921 CTCTACTGGGAGAGGTTGGAAACAGAGGAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 2033 CTCTACTGGGAGAGGTTGGAAACAGAGGAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2092
Qy 1981 CGTGGAGATCAAGCTTGGGACGCTCGGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 2093 CGTGGAGATCAAGCTTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2152
Qy 2041 GAACTGACCCAGTTAGTCTCTGCGAGGCTGTAAGGTTAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 2153 GAACTGACCCAGTTAGTCTCTGCGAGGCTGTAAGGTTAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2212
Qy 2101 CAGAGGACGCTTACGCTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
Db 2213 CAGAGGACGCTTACGCTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2272
Qy 2161 CTGCTTATCAGAGGGGGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220

Db	2273	CTGCTATTCAGAGGGGGCCACCGCTGCTCTCTGCCCATTGTCACGAGAGTCAACACAGC	2332
Qy	2221	TCACCTTCCATTGGCTTGGTGGAGAGAGCAGACTTTTACCACTGTCAAGATTGTCAAC	2280
Db	2333	TCACCTTCCATTGGCTTGGTGGAGAGAGCAGACTTTTACCACTGTCAAGATTGTCAAC	2392
Qy	2281	TACACTGTACGCTTCGGCCCTTGGGGGCTCAGAGATGCTTCCCTGGTCACTACTATACC	2340
Db	2393	TACACTGTACGCTTCGGCCCTTGGGGGCTCAGAGATGCTTCCCTGGTCACTACTATACC	2452
Qy	2341	AGCTCTGGAGAGAGACATTCTCATTTGGCGCCCTGAAACCAATTTACCAAGTACAGATTGG	2400
Db	2453	AGCTCTGGAGAGAGACATTCTCATTTGGCGCCCTGAAACCAATTTACCAAGTACAGATTGG	2512
Qy	2401	GTACAGTCCACGAGAGTATGATGAGTGGCCCTTTGGCTCCGTGCTAGAACGCTCCACC	2460
Db	2513	GTACAGTCCACGAGAGTATGATGAGTGGCCCTTTGGCTCCGTGCTAGAACGCTCCACC	2572
Qy	2461	GTGCTGACCGGCTTCAACACTCTCTGACCTGGCGCTGAGCCCTGACACCATCC	2520
Db	2573	GTGCTGACCGGCTTCAACACTCTCTGACCTGGCGCTGAGCCCTGACACCATCC	2632
Qy	2521	ACGTTTGGTTCACCTGGTGTCTCCCGACGAGCCCAATGGTGAGATTGTGGAGTATCTA	2580
Db	2633	ACGTTTGGTTCACCTGGTGTCTCCCGACGAGCCCAATGGTGAGATTGTGGAGTATCTA	2692
Qy	2581	ATTCTCTACAGCAACACACACCCAGCCGACCACTGAGTGGACACTGCTCCACACAGAG	2640
Db	2693	ATTCTCTACAGCAACACACACCCAGCCGACCACTGAGTGGACACTGCTCCACACAGAG	2752
Qy	2641	GGAAACATCTTCAGTGCAGAGTGCATGCGCTAGAGAGTGCACCTGGTATTTCTTCAG	2700
Db	2753	GGAAACATCTTCAGTGCAGAGTGCATGCGCTAGAGAGTGCACCTGGTATTTCTTCAG	2812
Qy	2701	ATGGGAGCCGACACAGAGTGGGCGCTGGCGCTTTTCCGCTTGCAGATGATGATTACT	2760
Db	2813	ATGGGAGCCGACACAGAGTGGGCGCTGGCGCTTTTCCGCTTGCAGATGATGATTACT	2872
Qy	2761	CTGCAAGAGACATTTCTCAGACTCTTTGGATGTGCAC	2796
Db	2873	CTGCAAGAGACATTTCTCAGACTCTTTGGATGTGCAC	2908
RESULT 6	AB052621		
LOCUS	Mus musculus DDM36E mRNA, complete cds.	6222 bp	linear
DEFINITION	Mus musculus DDM36E mRNA, complete cds.		
ACCESSION	AB052621		
VERSION	AB052621.1	GI:11862940	
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (sites)		
AUTHORS	Murakami, H., Nakamata, T., Nakayama, T., Yamamoto, H., Hosaka, T., Aoyama, T., Nagayama, S., Oka, M., Kiyono, T., Sasaki, M. S., Nakamura, T. and Toguchida, J.		
TITLE	Up-regulation of a ras effector and down-regulation of a cell adhesion molecule are associated with transformation of osteoblasts		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 6222)		
AUTHORS	Toguchida, J., Nakamata, T., Murakami, H., Nakayama, T. and Nakamura, T.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-DEC-2000) Junya Toguchida, Kyoto University, Institute for Frontier Medical Sciences, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto, Kyoto 606-8507, Japan (E-mail: togjun@frontier.kyoto-u.ac.jp, Tel: 81-75-751-4134, Fax: 81-75-751-4144)		
FEATURES	Location/Qualifiers		
source	1..6222		
	/organism="Mus musculus"		

gene	/mol_type="mRNA" /strain="C57BL/6BA" /db_xref="taxon:10090" /clone="DDM36E" /cell_line="MMC2" /cell_type="osteoblast" /note="similar to Nope (NM_020043)-vector:pDDM36EORP" 1..6222		
CDS	/gene="DDM36E" 50..3811 /gene="DDM36E" /note="similar to Nope (NP_064427)" /codon_start=1 /evidence=experimental /product="DDM36E" /protein_id="BAB19279.1" /db_xref="GI:11862941" /translation="MARADTGGTLLVTFCLLSARGBLPLPQSTTVKLSCDEGPLOVILGPQAVLDTLGTAGATGAPRPTVWSKGDVLEHNLHLPLNGSLWLSPLRQEDS DDEALRIWKVTEGSLASHPLGWSQVAVKLATLEBDFSLHPSQI VEENGTAR FECHTKGLPAPITWEXDOVTVPESRLITLPLNGVLOILDVQSDAGSYRCVNTSAR ORFQESALTVALLGSLSEATRGQDVVIYAAPENTTVSCQSVMEVCVASADPTFVSW VRQDKPISDVI VLGRNLILIASAQRHSGVYVCRANKPRTRDPATAAELRLAALP AISOAPEALSTRASTARFVCRASGPRPALHLDGIPLRPNRGRVKVGGGSLVIT QIGLQDAGTYQCVAAENSGARGTACAAPLAVVREGLFSAPTRTVATPLSSSVLVAMER PELHSEQIIGFSLHYOKARGVDVYQFVAVNDITLQVDRLEPNTDLYFVLYASOL CASRSTSPALVHTLDDVPSAAPQLTSSPNPBDIRVAMLPSSLSNGOVLKYKLEVG LGKESDQVFSTEVPGNETQLTNSLQPNKVYRIRISAGTCAGTGVPSQMMHRTIPGVH NQSVFPAPASLKVRAKMSLWSVQPPHPPTQISGYKLYREVTGTEBEADGDRPPGG RGDQAWDGVRLKVKYBELTQLVPGRLYEVKLVAFPKHEGDYAAVWKGKTEKAPT PDLPIQRGPPLPAHVAHESNSTSIWLRMKKPDFTTKVINYTVRPGPWLGRNASLV TYTSSGSDIILIGLKPFTKVFAYQSHGVDMGPGSVVERSTLPDRPSTPPSLRL SLPSTPVRHWCPPTPEPBGIVELIYLSNNHTQPEHQWTLITTEGSLPSEAHVGLLE SDTRYFFPMGAGTVEGPFGRLOQVITLQKTFSDSLDVHATVGLIVGCLGLCLLA CMGALRKSRRREALPGLSSGTFCNPALYTRARLGPSPVPAHSELSLHVRPQDWS PPSDVEDKAVHSLMGSDVSDCRHSKRKISWAQAGPNVAGSAGCGLPQSGSPRP ALTRALLPAGTGOTLLIQLVDAIKSNGRKKSPACRNOVEARVIVHSDFGASKGC PDLHLQDLEPEPELTAEPLTSGADVLSQADMGLRGLGSCOPTTSGPERLTCLPEA ASASCSDDLQSPSTAIBAPKSKQPKALCPITVSPSLPRAVPSAQVP"		
	BASE COUNT	1321 a	1772 c 1653 g 1476 t
	ORIGIN		
	Query Match	98.8%;	Score 2762.2; DB 10; Length 6222;
	Best Local Similarity	99.4%;	Pred. No. 0;
	Matches 2783;	Conservative 0;	Mismatches 13; Indels 3; Gaps 1;
Qy	1	GGGGAGCTGCCATTGCCCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGAGACCCCTG	60
Db	113	GGGGAGCTGCCATTGCCCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGAGACCCCTG	172
Qy	61	CAAGTGATCTCGGGCCCTGAGCAGGCTGTGTGCTGGACTGCACCTTTGGGGGCTACAGCT	120
Db	173	CAAGTGATCTCGGGCCCTGAGCAGGCTGTGTGCTGGACTGCACCTTTGGGGGCTACAGCT	232
Qy	121	GCTGGGCTCCGACCCAGGGTGACATGGAGCAAGGATGGAGACACTGTACTAGAGCATGAG	180
Db	233	GCTGGGCTCCGACCCAGGGTGACATGGAGCAAGGATGGAGACACTGTACTAGAGCATGAG	292
Qy	181	AACTGCACTGCTACCCAAATGGCTCCCTGTGGCTGTCTCAACCCCTAGACGAAGAGAC	240
Db	293	AACTGCACTGCTACCCAAATGGCTCCCTGTGGCTGTCTCAACCCCTAGACGAAGAGAC	352
Qy	241	AGCGATGATGAGAAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGCTG	300
Db	353	AGCGATGATGAGAAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGCTG	412
Qy	301	GCCCAAGCCCGCTAGGAGTGTGCTGCGCCAGCCAGGTGTCTGTGTTCAAGCTTGCACACTC	360
Db	413	GCCCAAGCCCGCTAGGAGTGTGCTGCGCCAGCCAGGTGTCTGTGTTCAAGCTTGCACACTC	472
Qy	361	GAAGACTTCTCTTGCAACCCGAGTCCCAAGTGTGTGGAGGAGAAACGGGACAGCAAGCTTT	420
Db	473	GAAGACTTCTCTTGCAACCCGAGTCCCAAGTGTGTGGAGGAGAAACGGGACAGCAAGCTTT	532

Qy	421	GAATGCCACCAAGGGCTTTCCAGCCCCCATCTATTCTTGGGAAAAGGACCAAGTGACC	480
Db	533	GAATGCCACCAAGGGCTTTCCAGCCCCCATCTATTCTTGGGAAAAGGACCAAGTGACC	592
Qy	481	GTGCTGAGGAGCCCGGCTCATCACTCTTCCAGAGTGGCTCTCCAGATCCTAGATGTC	540
Db	593	GTGCTGAGGAGTCCCGGCTCATCACTCTTCCAGAGTGGCTCTCCAGATCCTAGATGTC	652
Qy	541	CAGGACAGTATGATCAGGCTCTCTACCGCTGCGTGCCACCAATTCAGCCCGCCCAACGATTC	600
Db	653	CAGGACAGTATGATCAGGCTCTCTACCGCTGCGTGCCACCAATTCAGCCCGCCCAACGATTC	712
Qy	601	AGCCAGAGGGCTCGCTCACTGTGGCCCTCAGAGGGCTCTTTGGAGGCTTACAGGGGGCAG	660
Db	713	AGCCAGAGGGCTCGCTCACTGTGGCCCTCAGAGGGCTCTTTGGAGGCTTACAGGGGGCAG	772
Qy	661	GATGTGTCTATTGGCCAGCCGCCAGAGAAACCAACGATAGTGTCTGGACAGATGTAGTG	720
Db	773	GATGTGTCTATTGGCCAGCCGCCAGAGAAACCAACGATAGTGTCTGGACAGATGTAGTG	832
Qy	721	ATGAGTGTGGGCTCTGCTGACCCCAACCCCTTTTGTGCTCGGCTCCGACAGATGGA	780
Db	833	ATGAGTGTGGGCTCTGCTGACCCCAACCCCTTTTGTGCTCGGCTCCGACAGATGGA	892
Qy	781	AAGCCTATCTCCAGGATGTCACTGTCTGGGCGGACCAATCTACTCATCGCCAGCGG	840
Db	893	AAGCCTATCTCCAGGATGTCACTGTCTGGGCGGACCAATCTACTCATCGCCAGCGG	952
Qy	841	CAGCTGTGGCACTCTGAGTCTATGTCTGGCGAGCCAAAGAGCCCTCAGCGTGACTTC	900
Db	953	CAGCTGTGGCACTCTGAGTCTATGTCTGGCGAGCCAAAGAGCCCTCAGCGTGACTTC	1012
Qy	901	GCACCTGGCTGTGAGTCCGAGTCTGCTGCCCGCCAGCACTCGCAGGCCACCGAG	960
Db	1013	GCACCTGGCTGTGAGTCCGAGTCTGCTGCCCGCCAGCACTCGCAGGCCACCGAG	1072
Qy	961	CGCTCTCTCGGAGCGCGGGCCAGACACCGCGCTTCTGTGCGCGGGCTCGCGGGAGCCA	1020
Db	1073	CGCTCTCTCGGAGCGCGGGCCAGACACCGCGCTTCTGTGCGCGGGCTCGCGGGAGCCA	1132
Qy	1021	CGGCTCGGCTGTGCTGCAAGAGGATCCGTTGCGACCAATGGGCGCGTCAAG	1080
Db	1133	CGGCTCGGCTGTGCTGCAAGAGGATCCGTTGCGACCAATGGGCGCGTCAAG	1192
Qy	1081	GTGAGGGCGGTGGCGAGCTTGGTCACTCATCATCATCATCATCATCATCATCATCAT	1140
Db	1193	GTGAGGGCGGTGGCGAGCTTGGTCACTCATCATCATCATCATCATCATCATCATCAT	1252
Qy	1141	TACAGTGTGCTAGCAGAAAACAGCGCGGGAACCTGCTGTGCGCTGCGCCCTGGCGGTA	1200
Db	1253	TACAGTGTGCTAGCAGAAAACAGCGCGGGAACCTGCTGTGCGCTGCGCCCTGGCGGTA	1312
Qy	1201	GTGTGCGCAGGGGCTGCCAGCGCCCGACTCGGGTCAAGCCACCGCCCTGAGCAGC	1260
Db	1313	GTGTGCGCAGGGGCTGCCAGCGCCCGACTCGGGTCAAGCCACCGCCCTGAGCAGC	1372
Qy	1261	TCCTCTGTGCTGTGGCTGTGGAGCGGCTGAGTGTGACAGCGAGCAATCATTTGGCTTC	1320
Db	1373	TCCTCTGTGCTGTGGCTGTGGAGCGGCTGTGAGTTGCAAGCGAGCAATCATTTGGCTTC	1432
Qy	1321	TCTCTTCACTAACAAAGGCAAGGGAGTGACCAATGTGGAGTACCACTTTTGGCAGTAAAC	1380
Db	1433	TCTCTTCACTAACAAAGGCAAGGGAGTGACCAATGTGGAGTACCACTTTTGGCAGTAAAC	1492
Qy	1381	AATGACCAACAGAGCTGACGTTTGGGACCTTGGAAACCAACAGCGATTTAGTTCTTAC	1440
Db	1493	AATGACCAACAGAGCTGACGTTTGGGACCTTGGAAACCAACAGCGATTTAGTTCTTAC	1552
Qy	1441	GTGTGGGCTTACTCCAGCTGGGGCCAGCGGAACTCTCAGCCAGCCCTGGTGCATACA	1500
Db	1553	GTGTGGGCTTACTCCAGCTGGGGCCAGCGGAACTCTCAGCCAGCCCTGGTGCATACA	1612

Qy	1501	CTGAGCAGATGTCCCAAGCGCAGCACCCCAGCTTACCTTGTCCAGCCCCCAACCCCTCGGAC	1560
Db	1613	CTGAGCAGATGTCCCAAGCGCAGCACCCCAGCTTACCTTGTCCAGCCCCCAACCCCTCGGAC	1672
Qy	1561	ATCAGGGTGACATGGCTGCCCCCTCCAGCTGAGCAATGGACAGGTGCTGTGAGTAC	1620
Db	1673	ATCAGGGTGACATGGCTGCCCCCTCCAGCTGAGCAATGGACAGGTGCTGTGAGTAC	1732
Qy	1621	AAGATAGATGACGTTTGGGGAAAG--GAAGATCAGGTTTCTCCACCGAGGTGCCTGGA	1677
Db	1733	AAGATAGATGACGTTTGGGGAAAG--GAAGATCAGGTTTCTCCACCGAGGTGCCTGGA	1792
Qy	1678	AATGAGACACAATTAAGTTAAATCACTTTCAGCCAAACAAAGTGTACCGAGTCCGATTT	1737
Db	1793	AATGAGACACAATTAAGTTAAATCACTTTCAGCCAAACAAAGTGTACCGAGTCCGATTT	1852
Qy	1738	TCAGCTGGCACTGGCGTGGCTATGAGTTCCTTCTCAGTGGATGCAAGCAGGACACCT	1797
Db	1853	TCAGCTGGCACTGGCGTGGCTATGAGTTCCTTCTCAGTGGATGCAAGCAGGACACCT	1912
Qy	1798	GGTGTGCAACAACAGAGCCATGTTCCCTTTGGCCCTGACAGAAATGAAGGTGAGGGCAAG	1857
Db	1913	GGTGTGCAACAACAGAGCCATGTTCCCTTTGGCCCTGACAGAAATGAAGGTGAGGGCAAG	1972
Qy	1858	ATGAGTCTCTGTGTGTGTCAATGGCAGCCGCCCTTCAACCCCAACCAAGATCTCTGGATAC	1917
Db	1973	ATGAGTCTCTGTGTGTGTCAATGGCAGCCGCCCTTCAACCCCAACCAAGATCTCTGGATAC	2032
Qy	1918	AAACTCTACTTGGGAGAGGTGGGAAACAGAGAGGAGGAGGAGATGGTGTGACCGCCCCCAGGG	1977
Db	2033	AAACTCTACTTGGGAGAGGTGGGAAACAGAGAGGAGGAGGAGATGGTGTGACCGCCCCCAGGG	2092
Qy	1978	GGTGTGAGATCAAGCTTGGGAGCTCGGGCGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2037
Db	2093	GGTGTGAGATCAAGCTTGGGAGCTCGGGCGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2152
Qy	2038	TATGAACTGACCCAGTTAGTCTCTGGCAGCGCTGTACGAGGTGAAGCTCGTAGCTTTCAAC	2097
Db	2153	TATGAACTGACCCAGTTAGTCTCTGGCAGCGCTGTACGAGGTGAAGCTCGTAGCTTTCAAC	2212
Qy	2098	AAACAGAGAGCGGTACGCTGTGTGTGGAAGGGCAAGACGGAGAAAGGGCCCAACGCCA	2157
Db	2213	AAACAGAGAGCGGTACGCTGTGTGTGGAAGGGCAAGACGGAGAAAGGGCCCAACGCCA	2272
Qy	2158	GACCTGCTTATCCAGAGGGGGCCACCGCTGCTCTGCCCCATGTCCACGAGAGTCAAAC	2217
Db	2273	GACCTGCTTATCCAGAGGGGGCCACCGCTGCTCTGCCCCATGTCCACGAGAGTCAAAC	2332
Qy	2218	AGCTCCACTTTCATTTGGCTTTCGGTGGAAAGGACAGACTTTTACCACTGTCAAGATTTGTC	2277
Db	2333	AGCTCCACTTTCATTTGGCTTTCGGTGGAAAGGACAGACTTTTACCACTGTCAAGATTTGTC	2392
Qy	2278	AACCTACCTGTACGCTTTCGGCCCTCGGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT	2337
Db	2393	AACCTACCTGTACGCTTTCGGCCCTCGGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT	2452
Qy	2338	ACCAGCTCTGGAGAGACATTTCTATTGGGGCTGAAACCAATTTTACAGATGAGGTTT	2397
Db	2453	ACCAGCTCTGGAGAGACATTTCTATTGGGGCTGAAACCAATTTTACAGATGAGGTTT	2512
Qy	2398	GGGTACAGTCCCAACGAGGTGGATATGGATGGGGCCCTTTTGGCTCGGTGAGAACGCTCC	2457
Db	2513	GGGTACAGTCCCAACGAGGTGGATATGGATGGGGCCCTTTTGGCTCGGTGAGAACGCTCC	2572
Qy	2458	ACCTGTGCTGACGGCCCTTCAACACTCTTCTGACCTGCGCTGAGCCCCCTGACACCA	2517
Db	2573	ACCTGTGCTGACGGCCCTTCAACACTCTTCTGACCTGCGCTGAGCCCCCTGACACCA	2632
Qy	2518	TCACCGTTTGGTTTACTGCTGCTCCCCCAACGGAGGCCAATGGTGTGATTTGGAGTAT	2577
Db	2633	TCACCGTTTGGTTTACTGCTGCTCCCCCAACGGAGGCCAATGGTGTGATTTGGAGTAT	2692
Qy	2578	CTAATTTCTTACAGCAACAACCAACCAACCGAGCCCGGAAACACAGGTGGACACTGCTCACCACA	2637

Db	762	ATGAGTGGGCTCTGTGACCCACCCCTTTTGTGTCTCTGGTCCGACAGGGGAT	821
Qy	775	-----	774
Db	822	CAGAGTAAGAGGAACATGGGGATTCAAACACAGCGGCTGTGATGACGCTTTTGACAGGGATGT	881
Qy	775	-----	774
Db	882	AGTGGGAACCTCGTAGTTGGAGAGGGGTGTAGAGGGGACAGTTCCAGCGTCCCTCCAGC	941
Qy	775	-----	774
Db	942	CTGCTCCTTCTCTACACAGATGGAAGCCCTATCTCCACGGATGTATCGTTCTGGGCC	1001
Qy	815	GGACCAATCTACTCATCGCCAGCGCGAGCTCTGGCACTCTGGAGTCTATGTCTGCGGAG	874
Db	1002	GGACCAATCTACTCATCGCCAGCGCGAGCTCTGGCACTCTGGAGTCTATGTCTGCGGAG	1061
Qy	875	CCAAACAAGCCCTCAACGGTGAATTCGCACTGGGCTGTGAGCTCCGAGTGTCTGCTG	934
Db	1062	CCAAACAAGCCCGCAGCGGTGATTTGGCCACTGGGCTGTGAGCTCCGAGTGTCTGCTG	1121
Qy	935	CCCAAGCCATCTCGGAGGCAACCGAGGCGCTCTCGCGACCGGGCCAGCACCGCGGCT	994
Db	1122	CCCAAGCCATCTCGGAGGCGCCCGAGGCGCTCTCGCGACCGGGCCAGCACCGCGGCT	1181
Qy	995	TCGTGTGCGGCGCTCGGGGAGCACCGGCCGCGCTGCACTGGCTGCGACGAGCGGATCC	1054
Db	1182	TCGTGTGCGGCGCTCGGGGAGCACCGGCCGCGCTGCACTGGCTGCGACGAGCGGATCC	1241
Qy	1055	CGTTGCGACCAATGGG-CGGGTCAAGGTGAGGGCGGTGGCGGAGCTTGGTCACT	1113
Db	1242	CGTTGCGACCAATGGGCGGCTCAAGGTGAGGGCGGTGGCGGAGCTTGGTCACT	1301
Qy	1114	CAGATCGGCTTGAGGAGCGCTGGCTATACAGGTGCTGAGCAGAAAAACAGCGCGGAACT	1173
Db	1302	CAGATCGGCTTGAGGAGCGCTGGCTATACAGGTGCTGAGCAGAAAAACAGCGCGGAACT	1361
Qy	1174	GCTGTGCGGCTGGCGGCTGGCGGTAGTGTGCGGAGGGGCTGCCAGCGGCCCGACT	1233
Db	1362	GCTGTGCGGCTGGCGGCTGGCGGTAGTGTGCGGAGGGGCTGCCAGCGGCCCGACT	1421
Qy	1234	CGGGTCAAGCAGCAGCGGCTGAGCAGCTCTCTGTGCTGGTGGCTGGGAGCGGCTGAG	1293
Db	1422	CGGGTCAAGCAGCAGCGGCTGAGCAGCTCTCTGTGCTGGTGGCTGGGAGCGGCTGAG	1481
Qy	1294	TTGCAACAGCAGCAAAATCTTGGCTTCTCTTCTCACTACAAAAAGCAAGGGAGTGGAC	1353
Db	1482	TTGCAACAGCAGCAAAATCTTGGCTTCTCTTCTCACTACAAAAAGCAAGGGAGTGGAC	1541
Qy	1354	AATGTGGAGTACCAGTTTGAGTAAAGATGACACACAGAGCTGCAGGTTGCGGAGCTG	1413
Db	1542	AATGTGGAGTACCAGTTTGAGTAAAGATGACACACAGAGCTGCAGGTTGCGGAGCTG	1601
Qy	1414	GAAACCAACAGGATATAGTCTTACGTGTGGCTACTCCAGCTGGGGGCGAGCGGA	1473
Db	1602	GAAACCAACAGGATATAGTCTTACGTGTGGCTACTCCAGCTGGGGGCGAGCGGA	1661
Qy	1474	ACCTCCAGCCAGGCTGGTGCATACCTGGAAGATGTCCCAGCGCAGCAGCAGCTT	1533
Db	1662	ACCTCCAGCCAGGCTGGTGCATACCTGGAAGATGTCCCAGCGCAGCAGCAGCTT	1721
Qy	1534	ACCTGTCCAGCCAGGCTGGGATCATAGGGTGGATGGCTGCCCTGCCCTCCAGC	1593
Db	1722	ACCTGTCCAGCCAGGCTGGGATCATAGGGTGGATGGCTGCCCTGCCCTCCAGC	1781
Qy	1594	CTGAGCATGACAGGCTGAGTACAGATAGAGTACGTTTGGGGAAG---GAGAT	1650
Db	1782	CTGAGCATGACAGGCTGAGTACAGATAGAGTACGTTTGGGGAAGGAAGAAGAT	1841
Qy	1651	CAGGTTTTCTCCAGCGAGGCTCGGAAATGAGACACAACTTACGTTAAATCTCACTTCAG	1710
Db	1842	CAGGTTTTCTCCAGCGAGGCTCGGAAATGAGACACAACTTACGTTAAATCTCACTTCAG	1901

Qy	1711	CCAAACAAGGTGTACCGAGTCCGATTTCACTGTGGCACTGGCGCTGTGTATGGAGTCCCT	1770
Db	1902	CCAAACAAGGTGTACCGAGTCCGATTTCACTGTGGCACTGGCGCTGTGTATGGAGTCCCT	1961
Qy	1771	TCTCAGTGGATGACGACAGGACACCTGGTGTGCACAACACAGAGCCATGTTCCTTTGCC	1830
Db	1962	TCTCAGTGGATGACGACAGGACACCTGGTGTGCACAACACAGAGCCATGTTCCTTTGCC	2021
Qy	1831	CCTGCAGAAATTGAAGGTGAGGCAAGATGAGTCCCTGGTGGTGTCACTGGCAGCCGCC	1890
Db	2022	CCTGCAGAAATTGAAGGTGAGGCAAGATGAGTCCCTGGTGGTGTCACTGGCAGCCGCC	2081
Qy	1891	CCTCACCCCAACAGATCTCTGGATACAAACTCTACTCTGGGAGAGGTGGGAACAGAGAG	1950
Db	2082	CCTCACCCCAACAGATCTCTGGATACAAACTCTACTCTGGGAGAGGTGGGAACAGAGAG	2141
Qy	1951	GAGGAGATGGTGAACCGCCCCCGAGGGGTCTGGAGATCAAGCTTGGGAGCTCGGGGCC	2010
Db	2142	GAGGAGATGGTGAACCGCCCCCGAGGGGTCTGGAGATCAAGCTTGGGAGCTCGGGGCC	2201
Qy	2011	GTGCGGCTGAAGAGAAAGTGAAGCAGTATGAACCTGACCTGACCTGCTGGCAGGCTG	2070
Db	2202	GTGCGGCTGAAGAGAAAGTGAAGCAGTATGAACCTGACCTGACCTGCTGGCAGGCTG	2261
Qy	2071	TACGAGTGAAGCTCGTAGCTTTCAAACAACACAGAGACCGCTACGCTGTGTGGAG	2130
Db	2262	TACGAGTGAAGCTCGTAGCTTTCAAACAACACAGAGACCGCTACGCTGTGTGGAG	2321
Qy	2131	GGCAAGAAGGAGGCGCCCAAGCAGAGCTGTGCTATCAAGAGGGGCGCACCGCTGCT	2190
Db	2322	GGCAAGAAGGAGGCGCCCAAGCAGAGCTGTGCTATCAAGAGGGGCGCACCGCTGCT	2381
Qy	2191	CCTGCCCATCTCCACGAGAGTCAAAACAGCTCCACTTCCATTTGGCTTGGTGGAGAG	2250
Db	2382	CCTGCCCATCTCCACGAGAGTCAAAACAGCTCCACTTCCATTTGGCTTGGTGGAGAG	2441
Qy	2251	CCAGACTTTTACCCTGTCAAGATTTGTCAACTACACTGTACGCTTGGCGCCCTGGGGGCTC	2310
Db	2442	CCAGACTTTTACCCTGTCAAGATTTGTCAACTACACTGTACGCTTGGCGCCCTGGGGGCTC	2501
Qy	2311	AGGAATCTTCTCCCTGGTCACTACTATACAGCTCTCGGAGAGACATTTCTCATTTGGGCGC	2370
Db	2502	AGGAATCTTCTCCCTGGTCACTACTATACAGCTCTCGGAGAGACATTTCTCATTTGGGCGC	2561
Qy	2371	CTGAAACCATTTACCAAGTACAGTTTGGGTTCGGTACAGTCCCAAGAGTGGATATGGATGGG	2430
Db	2562	CTGAAACCATTTACCAAGTACAGTTTGGGTTCGGTACAGTCCCAAGAGTGGATATGGATGGG	2621
Qy	2431	CCCTTTGGCTCCGTCGTAGAACGCTCCACCTGCTGACCGGCTTCAACACCTCTTCT	2490
Db	2622	CCCTTTGGCTCCGTCGTAGAACGCTCCACCTGCTGACCGGCTTCAACACCTCTTCT	2681
Qy	2491	GACCTGGCGCTTGAGCCCGCTTGACACCATCCACCGTTCGGTTTACACTGTGTGTCCCGCCACG	2550
Db	2682	GACCTGGCGCTTGAGCCCGCTTGACACCATCCACCGTTCGGTTTACACTGTGTGTCCCGCCACG	2741
Qy	2551	GAGCCCAATTTGGTGGATTTGGAGTATCTAATTTCTTACAGCAACACCAACCCAGGCC	2610
Db	2742	GAGCCCAATTTGGTGGATTTGGAGTATCTAATTTCTTACAGCAACACCAACCCAGGCC	2801
Qy	2611	GAAACAAGTGGACACTGTCTCACACAGAGGGGAACATCTTCAAGATGGAGCTCCATGGC	2670
Db	2802	GAAACAAGTGGACACTGTCTCACACAGAGGGGAACATCTTCAAGATGGAGCTCCATGGC	2861
Qy	2671	CTAGAGAGTGAACCTCGGTATTTCTTCAAGATGGAGCTCCGCAAGAGGTGGGGCTGGG	2730
Db	2862	CTAGAGAGTGAACCTCGGTATTTCTTCAAGATGGAGCTCCGCAAGAGGTGGGGCTGGG	2921
Qy	2731	CCCTTTTCCCGCTTGCAGGATGTGATTAATCTCTGCAAGAGACATTTCTCAGACTCTTGGAT	2790
Db	2922	CCCTTTTCCCGCTTGCAGGATGTGATTAATCTCTGCAAGAGACATTTCTCAGACTCTTGGAT	2981

[illegible]

Db	2317	GAGCTGACCAGCTAGTCTCCCTGCGCGGCTGTACGAGGTGAAGCTCGTGGCTTTTCAACAAA	2376
Qy	2101	CACGAGACGGCTACGCTGCTGTGTGNAAGGCGAAGACGAGAGGCGCCACCGCCAGAC	2160
Db	2377	CATGAGATGGCTATTGACGACAGTGTGNAAGGCGAAGCGAGAGAGGCGCCGACCCAGAC	2436
Qy	2161	CTGCTATTACAGAGGGGGCCACCGCTGCTCTCTGCGCCCATGTCCAACGAGAGTCAAAACGC	2220
Db	2437	ATGCCTATTACAGAGGGGACACCCCTGCTCTCAAGCCCACTCCATGCGGAATCAAAACGC	2496
Qy	2221	TCCACTTTCCATTTGGCTTCGGTGGAGAGCGCAGACTTTACCACTGTGCAGATGTGCAC	2280
Db	2497	TCCACATCCATCTGGCTTCGGTGGAAAAGCGCAGATTTCCACCAGTCAAGATTTGTCAC	2556
Qy	2281	TACACTGTAGCTTTGGCGCCCTGGGGGCTCAGGAATGCTTTCCCTGGTGCACTACTATACC	2340
Db	2557	TACACTGTGGCTTCAGCCCCCTGGGGGCTCAGGATGGCTTCCCTGGTCACTATTACCC	2616
Qy	2341	AGCTCTGGAGAAGACATTTCTCATTTGGCGGCTGAAACCAATTATCAAGTACGAGTTTGGC	2400
Db	2617	AGTTCTGGAGAGACATCTCTCATTTGGCGGCTTGAAGCCAATTCCACAAATACGAGTTTGCA	2676
Qy	2401	GTACAGTCCCAACGAGTGGATATGATGGAGGCGCTTTGGGCTCCGTGTAGAAACGCTCCACC	2460
Db	2677	GTGCAGTCTCAACGCGCGTGGACATGGATGGGCGCTTTGGGCTCTGTGTGTGGAGCGCTCCACC	2736
Qy	2461	CTGCGTCAACGGGCTTCAACACCTCTCTTCGACTCGCTGCGCTGAGGCCCTCTGACACCATCC	2520
Db	2737	CTGCGTCAACGGGCTTCAACACCCCTTCGACTCGCTGCGCTGAGGCCCTCTGACACCGTCC	2796
Qy	2521	ACGGTTCCGGTTTACACTGGTGTCCCCCAACGGAGCCCAATGGTGAAGATTTGGAGTATCTA	2580
Db	2797	ACGGTTCCGGCTGCATGGTGTCCCCCAACGAGCCCAACGGGGAGATCGTGGAGTATCTG	2856
Qy	2581	ATTCTCTACAGCAACAACCAACACCGAGCCCGAACAACAGTGGACACTGTCTCAACCAAGAG	2640
Db	2857	ATCTGTACAGCAGCAACCAACACGACGCTTGAGCAACCAAGTGGACCTTTGCTCACCCACGAG	2916
Qy	2641	GGAAACATCTTCAGTGCAGAGGTCATGGGCTAGAGATGACACTCGGTATTTCTTCAAG	2700
Db	2917	GGAAACATCTTCAGTGTGAGGTCATGGGCTTGAGAGGACACTCGGTATTTCTTCAAG	2976
Qy	2701	ATGGGAGCCCGCAACAGAGGTGGGGCTTGGGGCCCTTTTCCGCTTGCAGGATGTGATTACT	2760
Db	2977	ATGGGAGCCCGCAACAGAGGTGGGACCTTGGGCTTCTTCCGCTTGAGGATGTGATTACT	3036
Qy	2761	CTGCAAGAGACATTTCTCAGACTCTCTTGGATGTGCAC	2796
Db	3037	CTCCAGAGAGAGCTGTCAAGCTCGCTGGAATGTGCAC	3072

RESULT 9

AX418406

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISMS

REFERENCES

AUTHOR:

1

TITLE

JOURNAL

00000000

FEATURES

AX418406 3741 bp DNA linear PAT 18-JUN-2002
Sequence 1 from Patent WO0206329.

AX418406
AX418406.1 GI:21523336

unidentified

unidentified
unclassified.

¹ Rastelli, L., Shimkets, R.A., Zerhusen, B., Malyankar, U.M. and

Padigaru, M.
Human polynucleotides and polypeptides encoded thereby

Patent: WO 0206329-A 1 24-JAN-2002;
Curagen Corporation (US)

Location/Qualifiers	
1 3241	

1. 3/41

```
/organism="unidentified"
/total_time="genomic DNA"
```

```
/mol_type="genomic DNA"  
_type="genomic DNA"  
_type="genomic DNA"
```

```

/ab_xref="taxon:32644"
/note="Nucleic acid with homology to NABP-like

```

NOVEL NOVEL ACID WITH NOVEL-LIKE

Db	2104	TATGAGCTGACCCAGCTAGTCCCTCGCGCGCTGTACAGGTGAAGCTCGTGGCTTTTCAC	2163
Qy	2098	AAACACAGGACGCGTACGCTGCTGTGTGGAAGGCGCAAGACGGAGAGGCGCCACGCCA	2157
Db	2164	AAACATGAGGATGCTATGACGACGTGTGGAAGGCGCAAGACGGAGAGGCGCCGACCA	2223
Qy	2158	GACCTGCTATCCAGAGGGGCGCACCGCTGCTCTCTGCCCATGTGCCACGAGTCAAC	2217
Db	2224	GACATGCTATCCAGAGGGGACCAACCCCTGCTCTCAGCCCAACGTCCTCATGCGGAATCAAAC	2283
Qy	2218	AGCTCCACTTCCATTTTGGCTTTCGGTGAAGACGACACTTTTACCACTGTCAAGATTGTC	2277
Db	2284	AGCTCCACATCCATCTGGCTTCGGTGAAGAACGACATTTTCACACAGTCAAGATTGTC	2343
Qy	2278	AACATCACTGTACGCTTCGCGCCCTCGGGGCTCAGGAAATGCTTCCCTGGTCACTACTAT	2337
Db	2344	AACATCACTGTGCGCTTCAGCCCTCGGGGCTCAGGAAATGCTTCCCTGGTCACTACTAT	2402
Qy	2338	ACGAGCTCTGGAGAGACATCTCATTTGGCGGCTGAAACATTTTACAAAGTACGAGTTT	2397
Db	2403	--CAGTTCTCGAGAGACATCTCTCATTTGGCGGCTTGAAGGCAATTCACAAATACGAGTTT	2460
Qy	2398	GCGGTACAGTCCCAACGAGTGGATATGATGGGGCCCTTTGGCTCGCTGCTGAGAACGCTCC	2457
Db	2461	GCACTGAGTCTCAACGCGTGGACATGATGGGCTTTTCGGCTCTGTGGTGGAGCGCTCC	2520
Qy	2458	ACCTGCTGACCGGCTTCAACACCTCTCTTCGACCTGCGCTGAGCCCTCGACACCA	2517
Db	2521	ACCTGCTGACCGTCTCTTCACACCCCATCGACCTGCGACTGAGCCCTCGACACCG	2580
Qy	2518	TCCAGCGTTTCGGTTACACTGGTGTCCCCCAACGAGAGCCCAATGGTGGAGTAT	2577
Db	2581	TCCAGCGTTTCGGTGCACATGGTGTCCCCCAACGAGAGCCCAACGAGAGTAT	2640
Qy	2578	CTAATCTCTACAGCAACACACACCGCGCGGACACACAGTGGACACTGCTCACAC	2636
Db	2641	CTGATCTCTACAGCAACACACACCGCGCGGACACACAGTGGACACTGCTCACAC	2700
Qy	2637	-----AGAGGGAACATCTTCAGTGCAGAGGTCATGGCTCAGAGTGCACCTCGGTAT	2691
Db	2701	CAGGTTGAGGGAACATCTTCAGTGCAGAGTTCATGGCTCAGAGTGCACCTCGGTAT	2760
Qy	2692	TTCTTCAAGATGGAGCGCCGACAGAGGTGGGCTGCGGCTTTTCCCGCTTGCAGGAT	2751
Db	2761	TTCTTCAAGATGGGCGCGACAGAGGTGGGCTGCGGCTTTTCCCGCTTGCAGGAT	2820
Qy	2752	GTGATCTCTGCAAGAGACATCTCTGACACTCTTGGATGTGCAC	2796
Db	2821	GTGATCAGCTCCAGGAGAGCTGTCTGACACTCTGCTGGACATGAC	2865
RESULT 10			
LOCUS	AB046848	5549 bp	mRNA linear PRI 22-FEB-2001
DEFINITION	Homo sapiens mRNA for KIAA1628 protein, partial cds.		
ACCESSION	AB046848		
VERSION	AB046848.1	GI:10047332	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Nagase, T., Kikuno, R., Nakayama, M., Hirose, M., and Ohara, O.		
TITLE	Prediction of the coding sequences of unidentified human genes. XVIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro		
JOURNAL	DNA Res.	7 (4),	273-281 (2000)
MEDLINE	20450683		
PUBMED	10997877		
REFERENCE	2 (bases 1 to 5549)		
AUTHORS	Ohara, O., Nagase, T. and Kikuno, R.		
TITLE	Direct Submission		

JOURNAL	Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdna@kazusa.or.jp, URL:htp://www.kazusa.or.jp/huge, Tel:81-438-52-3913, Fax:81-438-52-3914)		
FEATURES	Location/Qualifiers		
source	1. .5549		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="fh14247"		
	/notes="vector:pBluescriptII SK plus"		
gene	1. .5549		
	/genes="KIAA1628"		
CDS	<81..3023		
	/genes="KIAA1628"		
	/notes="Start codon is not identified. fh14247 cDNA clone for KIAA1628 has a 40-bp different sequence replacing a 108-bp sequence at the position between 737 and 846, and an 88-bp insertion after the position 983 of the sequence of KIAA1628, respectively."		
	/codon_start=1		
	/product="KIAA1628 protein"		
	/protein_id="BAB13454.1"		
	/db_xref="GI:10047333"		
	/translation="MGRKGRSALSSLSLERNQGEVKVIGRGAREGGGQAWGTGGGGRDLSGIPRSAPLAPAIQTQAPRALSTRASTARFVCRASGPRPALRWLNGAPLRNGRVKVOGGSLVITQIGLDAGYOCVAENSAGMACASLAVVREGLPSAPRTVATPLSSAVLWAMERPMHSEQLIGSLHYQKARGMDNVEYQFVANDTTELQVRDLPTWDYFYVAYSQLGASRTSTPALVHLLDDVPSAQLSPSPSNDIRVAWLPPLPSLNGQVYKIKYIGLGEKQIIFSTEVGRNETMLNSLQPNKPVRIISGKTAAGFGAPSQWMHRTTPSMHNSHPVFPAPALVKQAKMELVLSWMPPTPHQTSYKLYVREVGAREAGNDRLPGCRGDQAWDVGVPRLKKVKQVLTQLVGPLYKLVAFNKHEDGYAAVWKTEKAPAPMPLOQGPPLPAHVHAESENSTSLWLSWKPDFTTVKLNVTVRSPWGLNLSLVYTTSSGEDIILGGLKPTKYKPYAVOSHGVHMDGPRGSVVERSLPDRPSTPPSDRLSLSTPSTRLHWCPTTPEGEIVYILYSSNHTPEHMTSLTQGNIFSAEVHGLSDTRYPFMGARTTEVGPFSRLQDVTITLQEKLSLDLSPWAGHLSLPHPHQDPSPSPVDEBRAVHSLMGGVSGRSHSKRKISWAQPSGSSMAWACBELPOAGPRPALTRALLPAGTGTQLLQALVYDAIKGNRKKSPACRNOVEARVIVHSDPSASNGNDLHLQDLEPDEPLPPEAPDLISGVGDPGGGAWLDRHLGCGLAAGPDRILTLCLPEAASASCYPDLPQPEVLSETPGDSCQLKSPCLGASPGLP RSPVSSSA"		
BASE COUNT	1157 a	1546 c	1496 g 1350 t
ORIGIN			
Query Match	53.3%	Score	1491.4; DB 9; Length 5549;
Best Local Similarity	87.4%	Pred. No.	3.3e-298;
Matches	1633;	Conservative	0; Mismatches 236; Indels 0; Gaps 0;
Qy	928	CTTGTGCTGCCCCAGCCATCTCGCAGGCAACCCGAGGCGCTCTCGCGGAGCGGGCGCAGCAC	987
Db	264	CTNAGCGCTCCCGCCATCACTCAGGCGCCCGAGCGCTGTTCGCGGAGCGGGCGGAGACA	323
Qy	988	GCGCGCTTCTGTGTGCGGGCGTCCGGGAGCCACCGCCCGCGCTGCACTGCGCTGCACGAC	1047
Db	324	GCGCGCTTCTGTGTGCGCGCTCGGGGAGCGCGCGCTGCGCTGCGCTGCACAC	383
Qy	1048	GCGATCCGCTTGGACCCCAATGGCGCGCTCAGGTCAGGCGGCTGCGCGCAGCTTGCTC	1107
Db	384	GCGCGCGCTTGGCGCCCAACCGGCGCGCTCAAGGTCAGGCGCGCGCTGCGCGCTGCTC	443
Qy	1108	ATCAGCTCAGATCGGCTGCAGGACGCTGCTACTACAGTGCCTAGCAGAAACACGCGG	1167
Db	444	ATCAGCTCAGATCGGCTGCAGGACGCGCTACTACAGTGCCTGCTGAGAACACGCGG	503
Qy	1168	GGAATGCTGCTGCTGCGCTGCGCGCTGAGTGTGCGGAGGGGCTGCCAGCGCC	1227
Db	504	GGAATGCTGCTGCTGCGCGCTGCGCGCTGAGTGTGCGGAGGGGCTGCCAGCGCC	563
Qy	1228	CCGAGCTGGGTACAGCCACGCGCTGAGCAGCTCTCTGTGTGCTGCTGAGGAGCGG	1287
Db	564	CCGAGCTGGGTACAGCCACGCGCTGAGCAGCTCTCTGTGTGCTGCTGAGGAGCGG	623


```

repeat_region 11884 11922
/rpt_family=" (TC)n"
repeat_region 11930 11966
/rpt_family=" (TC)n"
repeat_region 11967 11994
complement (11967)
/rpt_family="B4"
repeat_region 12127
complement (12127)
/rpt_family="B1 MM"
repeat_region 12143
complement (12143)
/rpt_family="B4"
repeat_region 12252
complement (12170)
/rpt_family="RSINE1"
repeat_region 1253 1293
complement (1253)
/rpt_family="B1 MM"
repeat_region 13367
complement (13288)
/rpt_family="MIR"

Query Match 14.8%; Score 415; DB 10; Length 116580;
Best Local Similarity 97.7%; Pred. No. 1e-75;
Matches 421; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 931 GCTGCCCGACCATCTGCGAGGACCGAGCGCTCTCGCGAGCGCGGCGCACCGCG 990
Db 10502 GCTGCCCGACCATCTGCGAGGCGCGCGAGCGCTCTCGCGAGCGCGGCGCACCGCG 10443

QY 991 CGCTTGCTGTCGGCGGTCGCGGAGCCACGCGCCCGCTGCACTGGCTGCAAGCGGG 1050
Db 10442 CGCTTGCTGTCGGCGGTCGCGGAGCCACGCGCCCGCTGCACTGGCTGCAAGCGGG 10383

QY 1051 ATCCGCTGCGACCCCAATGCGCGCGCTCAAGGTCGCGGCGGTCGCGAGCTTGGTCATC 1110
Db 10382 ATCCGCTGCGACCCCAATGCGCGCGCTCAAGGTCGCGGCGGTCGCGAGCTTGGTCATC 10323

QY 1111 ACTCAGATCGGCTCGAGGACGCTGCTACTACAGTGCAGGCGGTCGAGCAAAACAGCGCGGA 1170
Db 10322 ACTCAGATCGGCTCGAGGACGCTGCTACTACAGTGCAGGCGGTCGAGCAAAACAGCGCGGA 10263

QY 1171 ACTGCTGTCGCGTGGCGCCCTCGCGGTAGTGGTGGCGAGGGTGCACGCGCCCG 1230
Db 10262 ACTGCTGTCGCGTGGCGCCCTCGCGGTAGTGGTGGCGAGGGTGCACGCGCCCG 10203

QY 1231 ACTCGGTGACAGCGCGCGCTGAGCAGCTCTCTGCTGCTGCTGGCTGGAGCGGCT 1290
Db 10202 ACTCGGTGACAGCGCGCGCTGAGCAGCTCTCTGCTGCTGCTGGCTGGAGCGGCT 10143

QY 1291 GAGTTGCACAGCGAGCAAAATCATTTGGCTCTCTTCACTACCAAGGCAAGGCGGAGTG 1350
Db 10142 GAGTTGCACAGCGAGCAAAATCATTTGGCTCTCTTCACTACCAAGGCAAGGCGGAGTG 10083

QY 1351 GACATGTGGA 1361
Db 10082 TCTACTCTGA 10072

RESULT 12
AC110235/c
LOCUS AC110235 158055 bp DNA linear HTG 15-MAY-2002
DEFINITION HTG; Mus musculus clone RP23-100M12, WORKING DRAFT SEQUENCE, 17 ordered
pieces.
ACCESSION AC110235
VERSION AC110235.5 GI:20800205
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 158055)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-100M12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 158055)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgaiter,B.,

```

```

Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Chopel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Melrim,J., Meneus,L.,
Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Olivier,J., Peterson,K., Phunkhang,P., Riebeck,M., Rieck,R., Riese,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Stojanovic,N.,
Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 158055)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgaiter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Chopel,Y., Colangelo,M., Collins,S., Collamore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McKernan,K., Melrim,J., Meneus,L.,
Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Olivier,J., Peterson,K., Phunkhang,P., Riebeck,M., Rieck,R., Riese,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Stojanovic,N.,
Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (15-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 15, 2002 this sequence version replaced gi:20531900.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22383
Center clone name: 100_M_12
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 153807 bases at least Q40
Consensus quality: 155134 bases at least Q30
Consensus quality: 155784 bases at least Q20
Insert size: 218000; agarose-fp
Insert size: 156455; sum-of-ctnigs
Quality coverage: 7.1 in Q20 bases; agarose-fp
Quality coverage: 10.0 in Q20 bases; sum-of-ctnigs
-----

```

TITLE JOURNAL REFERENCE AUTHORS

TITLE JOURNAL COMMENT

```
/note="assembly_fragment"
```

misc_feature

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, J., Cleveland, C., Cockrell, R., Cox, C., Coyie, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, J., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackeon, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Labow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Loulesged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoletemeh, O., Okwuch, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfrankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shwartzbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanli, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczky, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 224227)
Worley, K.C.

Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 224227)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:22857495.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
. Center: Baylor College of Medicine

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: CH230-20218
Center clone name: GIBA
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 211258 bases at least Q40
Consensus quality: 214044 bases at least Q30
Consensus quality: 215992 bases at least Q20
Estimated insert size: 220399; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 221791: contig of 221791 bp in length
* 221792 221891: gap of unknown length
* 221892 222909: contig of 1018 bp in length
* 222910 223009: gap of unknown length
* 223010 224227: contig of 1218 bp in length.
Location/Qualifiers
1..224227
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-20218"
1..1752
/note="wgs end extension
clone end:Sp6"
3617_-3835
/note="clone boundary
clone end:Sp6"
site:ECORI
end:sequence:BH334357"
220120..221791
/note="wgs contig"
misc_feature
59284 a 52300 c 50849 g 55959 t 6835 others
BASE COUNT
ORIGIN
Query Match 13.5%; Score 376.8; DB 2; Length 224227;
Best Local Similarity 88.7%; Pred. No. 7,5e-68;
Matches 408; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 919 CTCGAGTGTCTGTCGCCCGACGCAATCTGCGAGGACCCGAGCGCTCTCGCGGACGCG 978
Db 109894 CTCCTGCTCTCAGCTGCCCGACGCAATCTGCGAGGCGCCGAGCGCTGTGCGGACG 109835
QY 979 GCGACACGCGCGCTTCTGTGCGGGGTCGCGGGGACCCAGCGCGCGCTGCGACTGG 1038
Db 109834 GCGACGCTGCGCGCTTCTGTGCGGGGTCGCGGGGACCCAGCGCTGCGACTGG 109775
QY 1039 CTGCGACGAGCGGATCCGCTTGTGCGAGCCCAATGCGCGCGCTCAAGGTGCGAGCGCGTGGCGGC 1098
Db 109774 CTGCGACGAGCGGCGCGCTTGTGCGAGCCCAATGCGCGCGCTCAAGGTGCGAGCGCGTGGCGGC 109715
QY 1099 AGCTTGGTCACTACATGATCGGCTGCGAGGACGCTGCTACTACAGTGCCTAGCAGAA 1158
Db 109714 AGCTTGGTCACTACATGATCGGCTGCGAGGACGCTGCTACTACAGTGCCTAGCAGAA 109655
QY 1159 AACAGGCGGGAACCTGCTGTCGCCCTGCGCGGCTAGTGTGCGCGGAGGCTG 1218
Db 109654 AACAGGCGGGAACCCCTGCTGTCGCCCTGCGCGGCTAGTGTGCGCGGAGGCTG 109595
QY 1219 CCCAGCGCCCGACTCGGCTCACAGCCCGCGCTCTCTGTGCTGTGGCC 1278

Quality coverage: 3.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 27 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1477:	contig of 1477 bp in length
*	1478	1577: gap of 100 bp
*	1578	2785: contig of 1208 bp in length
*	2786	2885: gap of 100 bp
*	2886	4859: contig of 1974 bp in length
*	4860	4959: gap of 100 bp
*	4960	6416: contig of 1457 bp in length
*	6417	6516: gap of 100 bp
*	6517	8739: contig of 2223 bp in length
*	8740	8839: gap of 100 bp
*	8840	10868: contig of 2029 bp in length
*	10869	10968: gap of 100 bp
*	10969	13737: contig of 2769 bp in length
*	13738	13837: gap of 100 bp
*	13838	15329: contig of 1492 bp in length
*	15330	15429: gap of 100 bp
*	15430	17938: contig of 2509 bp in length
*	17939	18038: gap of 100 bp
*	18039	19981: contig of 1943 bp in length
*	19982	20081: gap of 100 bp
*	20082	23587: contig of 3506 bp in length
*	23588	23687: gap of 100 bp
*	23688	27447: contig of 3760 bp in length
*	27448	27547: gap of 100 bp
*	27548	30161: contig of 2614 bp in length
*	30162	30261: gap of 100 bp
*	30262	34737: contig of 4476 bp in length
*	34738	34837: gap of 100 bp
*	34838	40101: contig of 5264 bp in length
*	40102	40201: gap of 100 bp
*	40202	40309: contig of 3838 bp in length
*	40400	44139: gap of 100 bp
*	44140	49950: contig of 5811 bp in length
*	49951	50050: gap of 100 bp
*	50051	55535: contig of 5505 bp in length
*	55556	55625: gap of 100 bp
*	55656	60496: contig of 4841 bp in length
*	60497	60596: gap of 100 bp
*	60597	68544: contig of 7948 bp in length
*	68545	68644: gap of 100 bp
*	68645	76940: contig of 8296 bp in length
*	76941	77040: gap of 100 bp
*	77041	85480: contig of 8440 bp in length
*	85481	85580: gap of 100 bp
*	85581	93905: contig of 8325 bp in length
*	93906	94005: gap of 100 bp
*	94006	108063: contig of 14058 bp in length
*	108064	108163: gap of 100 bp
*	108164	123496: contig of 15333 bp in length
*	123497	123596: gap of 100 bp
*	123597	139965: contig of 16369 bp in length
*	139966	140065: gap of 100 bp
*	140066	176282: contig of 36217 bp in length.

Location/Qualifiers

1. 176282

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP11-349H15"

/clone_lib="RPCI-11 Human Male BAC"

1. 1477

ture

```
RES
source

misc_f
```


Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, J., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 198295)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckigalier, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, J., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (26-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 26, 2002 this sequence version replaced gi:20148056.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://fcp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
 Center project name: L22971
 Center clone name: 349_G13

----- Location/Qualifiers
 1..198295
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-349G13"
 /clone_lib="RP11 Human Male BAC"
 complement(1)..79)
 /rpt_family="AluS"
 complement(183..494)
 /rpt_family="AluSg"
 complement(1216..1416)
 /rpt_family="L1MB8"
 complement(1417..1709)

FEATURES
source

repeat_region
 repeat_region
 repeat_region
 repeat_region

```
repeat_region 16409..16481
repeat_region /rpt family="MBR5A"
16482..16779
repeat_region /rpt family="AluSq"
16780..16874
repeat_region /rpt family="MBR5A"
complement(17247..17543)

Query Match      11.8%; Score 331.2; DB 9; Length 198295;
Best Local Similarity 80.6%; Pred. No. 2.1e-58; Indels 0; Gaps 0;
Matches 387; Conservative 0; Mismatches 93;

QY 867 CTGCCGAGCAACAAGCCCTCACGCGTGACTTCGCCACTGGCGGTGCTGAGCTCCGAGT 926
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 927 GCTTGTGCCCCCAGCCATCTGCGAGGCACCCGAGGCGCTCTCCGGAACGCGGGCCAGCAC 986
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89588 CCTAGCGGCTCCCGCCATCACTCAGCGGCCGAGGCGCTGTCCGGAACGCGGGCCAGCAC 89647
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 987 CGCGCGCTTCGTGTGCGGGCGCTCCGGGAGCCACGGCCGCGGCTGCACTGGCTGCACGA 1046
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89648 AGCGCGCTTCGTGTGCGGGCGCTCCGGGAGCCGCGGCGCTGCGCTGGCTGCACAA 89707
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1047 CGGGATCCCGTTGCGCAACCAATGGGCGCGTCAAGGTGCAGGCGGTGCGCGCAGCTTGGT 1106
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89708 CGGGGCGCGCTGCGGCGCCCAACGGGCGCGTCAAGGTCCAGGGCGGCGGTGGCAGCTGGT 89767
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1107 CATCACTCAGATCGGCTCGAGGACCGCTGGCTACTACCAAGTGGTAGCAGAAACAGCGC 1166
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89768 CATCACACAGATCGGCTCGAGGACCGCGCTACTACCAAGTGGTAGCAGAAACAGCGC 89827
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1167 GGGAACTGCTGTGCGCGTGGCCCTGGCGGTAGTGTGCGGAGGGGCTGCCAGCGC 1226
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89828 GGGAACTGCGGTGGCTGCGCGGTGGCTGGCGGTGGCTGGCGAGGGGCTGCCAGCGC 89887
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1227 CCGCACTCGGGTCACAGCCACGCGCTGAGCAGCTCTCTGTGCTGGTGGCTGGGAGCG 1286
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89888 CCCCACGCGGTCACCTGCTAGGCCACTGAGCAGCTCCGCTGTGTGGTGGCTGGGAGCG 89947
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1287 GCTGAGTTGCACAGCGAGCAATCATTTGGCTTCTCTTCACTACCAAAAGGCAAGGGG 1346
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89948 GCCCGAGATGCACAGCGAGCAGATCATCGGCTTCTCTCTCCACTACCAAGGCAAGGGG 90007
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: October 15, 2003, 04:35:54
Job time : 10466.9 secs